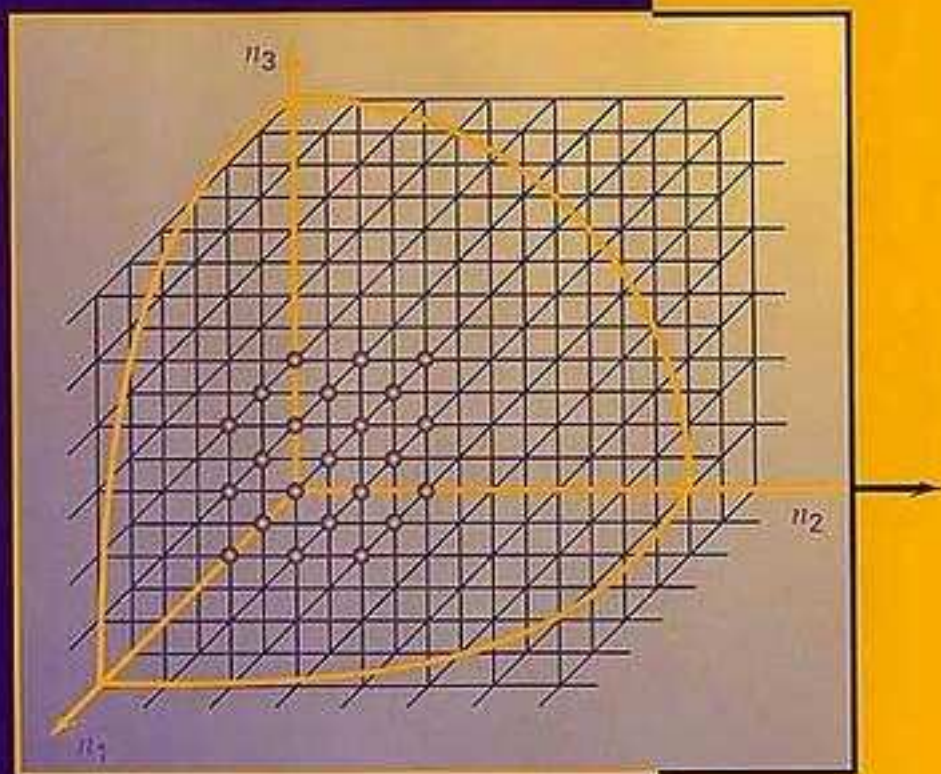


Quantum Physics

3rd edition



Stephen Gasiorowicz

www.ebook3000.com

Quantum Physics

Third Edition

Stephen Gasiorowicz

University of Minnesota



Quantum Physics III

Errata

p.21. In problem 2, replace ...use your estimate of the sun's surface temperature to... by and use 6000K as an estimate of the sun's surface temperature to...

p.40 In the solution to Example 2-4 (b) the second equation should have $\langle x^2 \rangle$ on the left side.

.64. In problem 11, the last equation should read

$$\psi(x,t) = A \left(\sin \frac{\pi x}{a} \right)^5$$

p.90. In problem 2, the definition of $V(x)$ in the third line should read

$$V(x) = 0 \quad a < x$$

p.92. In problem 2, on the right side of the two equations make the replacements

$$\tanh ab \rightarrow \tanh \alpha b; \quad \coth ab \rightarrow \coth \alpha b$$

p. 103 In equation (5-49), in lines 4 and 5 there should be a dx inserted.

p.118 In problem 7, the first line should read: Use the results of problems 5 and 6

In problem 16, the first equation should read

$$e^{\lambda A} f(A^+) |0\rangle = f(A^+ + \lambda) |0\rangle$$

p.119. In problem 17, the last term of the last equation should read $\frac{\lambda^2}{2!} [A, [A, A^+]]$

p.126 In equation (7-48) the denominator should be $(l+m)!$

p.148 In equation (9-21) the summation sign should read $\sum_{k,l}$

p.154. In the first un-numbered equation which lists the eigenvectors, the last one should be associated with $\lambda = -1$.

p.156. Problem 1: In the second line use :...left-hand 5 x 5 corner of the infinite array.)

The two equations referred to in problems 2 and 3 should be (6-36) instead of (6-4).

p.191, The material just below Figure 12-1 should read as follows:

It follows that the energy shift is

$$\Delta E = \frac{1}{4} m_e c^2 (Z\alpha)^4 \frac{\begin{Bmatrix} \ell \\ -\ell-1 \end{Bmatrix}}{n^3 l(l+1/2)(l+1)} \quad (12-16)$$

valid for $l \neq 0$. When the effects of H_1 and H_2 are combined, we get

$$\Delta E = -\frac{1}{2} m_e c^2 (Z\alpha)^4 \left(\frac{1}{j+1/2} - \frac{3}{4n} \right) \quad (12-17)$$

valid for both values of $l = j \pm 1/2$. It is necessary....

p.197. In problem 4. The final states in both cases should be $1^2S_{1/2}$. In part (b) the first wavelength should be 589.592 nm.

p.208 replace material from below Eq. (13-54) to below Eq. (13-57) by

Note that there is quite a lot of degeneracy in the problem: There are as many solutions for a given E as there are sets of integers $\{n_1, n_2, n_3\}$ that satisfy (13-54). Degeneracy is usually associated with the existence of mutually commuting observables and this example is no exception. In general when $H = H_x + H_y + H_z$ where

$$\begin{aligned} H_x &= \frac{p_x^2}{2m} + V_1(x) \\ H_y &= \frac{p_y^2}{2m} + V_2(y) \\ H_z &= \frac{p_z^2}{2m} + V_3(z) \end{aligned} \quad (13-55)$$

the eigenfunctions of H have the form

$$\psi_{E_1 E_2 E_3}(x, y, z) = u_{E_1}(x) v_{E_2}(y) w_{E_3}(z) \quad (13-56)$$

where $u_{E_1}(x), v_{E_2}(y), w_{E_3}(z)$ are eigenfunctions of H_x, H_y , and H_z respectively, with eigenvalues E_1, E_2 and E_3 , and

$$E = E_1 + E_2 + E_3 \quad (13-57)$$

p.238 In the expression for $c_n(\infty)$ the second line should read

$$\frac{qE}{i\hbar} \langle x | x | 0 \rangle \int_{-\infty}^{\infty} dt' e^{-i\omega t'} e^{-t'^2/\tau^2}$$

p.249 The two equations following (16-22) should read

$$\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial t} (e^{-i\Lambda} \psi') = -i \frac{\partial \Lambda}{\partial t} \psi + e^{-i\Lambda} \frac{\partial \psi'}{\partial t}$$

$$-i\hbar \nabla \psi = -i\hbar \nabla (e^{-i\Lambda} \psi') = -\hbar \nabla \psi - i\hbar e^{-i\Lambda} \nabla \psi'$$

p.255 In the line following Eq. (16-71) replace L_1, L_2 by the product $L_1 L_2$.

p.260 In the line above Eq. (17-3) replace t-me-independent by time-dependent

p. 264 Two lines below Eq. (17-30) replace $m = m_i$ by $m = -m_i$

p.266 In Eq. (17-37) the r.h.s. should read $\approx \frac{g\hbar}{4} \left| \frac{k}{p} \right| = \dots$

p.316. line 8 above Eq. (20-1) replace protons by photons.

530.12
G AS

Acquisitions Editor	<i>Stuart Johnson</i>
Marketing Manager	<i>Robert Smith</i>
Production Editor	<i>Barbara Russiello</i>
Senior Designer	<i>Karin Kincheloe</i>
Illustration Editor	<i>Sandra Rigby</i>
Cover Designer	<i>Carol C. Grobe</i>
Cover Photo	© <i>Digital Vision</i>

This book was set in 10/12 Times Roman by UG / GGS Information Services, Inc. and printed and bound by Hamilton Printing Company. The cover was printed by Phoenix Color Corporation.

This book is printed on acid free paper. This book is printed on acid free paper.

Copyright © 2003 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, (978) 750-8400, fax (978) 646-8600. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008.

To order books or for customer service please, call 1(800)-CALL-WILEY (225-5945).

Main Book ISBN 0-471-05700-2

Wiley International Edition ISBN 0-471-42945-7

Printed in the United States of America

10 9 8 7 6 5 4 3 2

91-33315

Preface

The first edition of *Quantum Physics* was published about thirty years ago. The guidelines that I set out in the preface to that edition as well as its general approach are ones to which I still subscribe. I wrote:

This book is intended to serve as an introduction to quantum physics. In writing it, I have kept several guidelines in mind:

1. First, it is helpful for the development of intuition in any new field of study to start with a base of detailed knowledge about simple systems. I have therefore worked out a number of problems in great detail, so that the insight thus obtained can be used for more complex systems.
2. Every aspect of quantum mechanics has been helpful in understanding some physical phenomenon. I have laid great stress on applications at every stage of development of the subject. Although no area of quantum mechanics is totally developed, my intention is to bridge the gap between a modern physics course and the more formal development of quantum mechanics. Thus, many applications are discussed, and I have stressed order-of-magnitude estimates and the importance of numbers.
3. In keeping with the level of the book, the mathematical structure has been kept as simple as possible. New concepts, such as operators, and new mathematical tools necessarily make their appearance. I have dealt with the former more by analogy than by precise definition, and I have minimized the use of new tools insofar as possible.

There were few changes in the second edition. This, the third edition, is not fundamentally different in spirit from the first two, but it differs in detail in many ways.

The principal changes are the following:

- (a) The ordering of chapters has been changed. This edition of the book divides more neatly into the basics of quantum mechanics and the most important applications that one would want to cover in a one-semester course versus a series of chapters on more general applications. Because I have tended to build new material on the base of previously discussed material, this change makes the book more useful for a one-semester course.
- (b) In response to advice from a number of users of the book, I have expanded somewhat the discussion of the general structure of wave mechanics and the more abstract description of quantum mechanics (including the Dirac notation). In keeping with guideline 3 listed above, I have avoided the introduction of a mini-course on linear algebra. This beautiful mathematical development is best treated in a first year graduate course.
- (c) I did not wish to increase the size of the book and have therefore shortened parts of the text by putting some material on the book's web site [www.wiley.com/college/gasiorowicz]. This material consists of several kinds of supplements:
 - some longer derivations or arguments that are not central to the development of a particular subject; examples include the discussion of the derivation of the blackbody radiation formula, as derived by Einstein, and topics in the addition of angular momentum

- material that has previously appeared in appendices, such as the WKB approximation and the derivation of the exponential decay formula for atoms
- applications that previously appeared in the text but usually had to be left out because of lack of time. Examples include the details of the calculation of barrier penetration in nuclei, solving the Schrödinger equation for periodic potentials, and some topics involving the radial equation for square wells.
- some new material, including a brief discussion of probability, the explicit solution of the differential equations for the spherical harmonics, and some topics related to the new Chapter 20
- In addition, I have shortened the introductory chapter, because most students will have had an introductory course in Chapter 1. This chapter has served as a review of the Old Quantum Theory. This material appears in every textbook on modern physics in one form or another. I did not eliminate the chapter because it still serves as a useful compendium of the workable rules developed in the decade or so before the creation of quantum mechanics, and it highlights the differences between classical and quantum physics.

Further changes involve the condensation of the treatment of helium, multielectron atoms, and molecules into a single chapter, with many details left to the web site. The treatment of an electron in a magnetic field has been moved to a much later place in the book so that three chapters containing this topic, radioactive decays, and selected topics on radiation form a unit (Chapters 16, 17, 18). The treatment of the Einstein A and B coefficients appears as a supplement to Chapter 1, because it provides a better introduction to the problem of blackbody radiation. The separate chapter on the calculation of the cross section for the photoelectric effect has been eliminated, with this material now incorporated into the chapter on collision theory as an example of an inelastic process.

(d) New topics include

- a discussion of the three-level system, with a brief mention of dark states and electromagnetically induced transparency.
- a new chapter that addresses issues on the foundations of quantum mechanics that have stimulated some beautiful experimental work. I do not deal in any detail with questions of the “interpretation of quantum mechanics” but concentrate on a small selection of experiments that illustrate in some measure just how counterintuitive quantum mechanics is when viewed through “classical” eyes.

ACKNOWLEDGMENTS

This edition was made possible by many people at Wiley who steered the manuscript through pages and finally into a bound book. Thanks to Stuart Johnson, my acquisitions editor, and to Robert Smith, marketing manager. Barbara Russiello, production editor, guided the book through composition. Helen Walden copyedited the manuscript. Sandra Rigby was the illustration editor for this edition. Karin Kincheloe designed the book and Carol C. Grobe designed the cover. Joyce Franzen at University Graphics made sure the pages came out correctly. My thanks to all who worked on this book.

The following reviewers helped in creating this third edition of *Quantum Physics*. I especially wish to thank Steven Pollock for all of his time on this manuscript.

Duncan Carlsmith, *University of Wisconsin-Madison*

Gabriel Karl, *University of Guelph*

Xinsheng Ling, *Brown University*

K.T. Mahanthappa, *University of Colorado-Boulder*

Charles Meitzler, *Sam Houston State University*

Steven Pollock, *University of Colorado-Boulder*

Chandralekha Singh, *University of Pittsburgh*

Mark Srednicki, *University of California-Santa Barbara*

Paul Ulmer, *Old Dominion University*

John VENABLES, *Arizona State*

Jens Zorn, *University of Michigan*

Stephen Gasiorowicz
Minneapolis, Minnesota

Table of Contents

1. The Emergence of Quantum Physics 1

1-1	Blackbody Radiation	1
1-2	The Photoelectric Effect	5
1-3	The Compton Effect	7
1-4	Wave Properties of Matter and Electron Diffraction	10
1-5	The Bohr Atom	15
	The Rutherford Planetary Model	15
	The Bohr Postulates	15
	The Correspondence Principle	20
	Problems	21

SUPPLEMENTS ON THE WEB SITE:

1-A	<i>Einstein's Approach to Planck's Law</i>	W-1
1-B	<i>Estimate of Lifetime of a Rutherford Atom</i>	W-4

2. Wave Particle Duality, Probability, and the Schrödinger Equation 23

2-1	Radiation as Particles, Electrons as Waves	23
2-2	Plane Waves and Wavepackets	25
	*How Wave Packets Move	27
2-3	The Probability Interpretation of the Wave Function	28
2-4	The Schrödinger Equation	30
2-5	The Heisenberg Uncertainty Relations	32
	Diffraction of a Photon Beam	32
	Inability to Localize Bohr Orbits	33
	Interim Summary	34
2-6	More on the Probability Interpretation	34
	The Importance of Phases	35
	The Probability Current	35
2-7	Expectation Values and the Momentum in Wave Mechanics	36
	The Momentum in Wave Mechanics	37
	Wave Function in Momentum Space	38
	Problems	41

SUPPLEMENTS ON THE WEB SITE:

2-A	<i>The Fourier Integral and Delta Functions</i>	W-6
2-B	<i>A Brief Tutorial on Probability</i>	W-11

3. Eigenvalues, Eigenfunctions, and the Expansion Postulate 44

3-1	The Time-Independent Schrödinger Equation	44
3-2	Eigenvalue Equations	45
3-3	The Eigenvalue Problem for a Particle in a Box	47
3-4	The Expansion Postulate and Its Physical Interpretation	52
	The Interpretation of the Expansion Coefficients	53

xii Table of Contents

3-5	Momentum Eigenfunctions and the Free Particle	56
	Normalization of the Free Particle Wave Function	57
	Degeneracy	59
3-6	Parity	60
	Problems	63
4.	One-Dimensional Potentials	66
4-1	The Potential Step	66
4-2	The Potential Well	69
4-3	The Potential Barrier	71
4-4	An Example of Tunneling	73
4-5	Bound States in a Potential Well	75
4-6	Deltas Function Potentials	81
4-7	The Harmonic Oscillator	85
	Problems	90
<i>SUPPLEMENTS ON THE WEB SITE:</i>		
4-A	<i>The Wentzel-Kramers-Brillouin-Jeffreys Approximation</i>	W-14
4-B	<i>Tunneling in Nuclear Physics</i>	W-16
4-C	<i>Periodic Potentials</i>	W-19
	<i>The Kronig-Penney Model</i>	W-21
5.	The General Structure of Wave Mechanics	94
5-1	Eigenfunctions and Eigenvalues	94
	The Hamiltonian Operator	94
5-2	Other Observables	96
	The Interpretation of the Expansion Coefficients	97
5-3	Vector Spaces and Operators	98
5-4	Degeneracy and Simultaneous Observables	100
5-5	Time Dependence and the Classical Limit	103
	Problems	105
<i>SUPPLEMENT ON THE WEB SITE:</i>		
5-A	<i>Uncertainty Relations</i>	W-26
6.	Operator Methods in Quantum Mechanics	107
6-1	Rephrasing Wave Mechanics—An Abstract View of Quantum Mechanics	107
	Projection Operators	111
6-2	The Energy Spectrum of the Harmonic Oscillator	112
6-3	From Operators Back to the Schrödinger Equation	114
6-4	The Time Dependence of Operators	116
	Problems	117
7.	Angular Momentum	120
7-1	The Angular Momentum Commutation Relations	121
7-2	Raising and Lowering Operators for Angular Momentum	122
7-3	Representation of $ l, m\rangle$ States in Spherical Coordinates	124
7-4	Comments on the Expansion Theorem	128
	Problems	129

SUPPLEMENTS ON THE WEB SITE:

- 7-A Rotational Invariance** W-28
Invariance Under Rotations About the z-Axis W-28
7-B Angular Momentum in Spherical Coordinates W-30

**8. The Schrödinger Equation in Three Dimensions
 and the Hydrogen Atom 131**

- 8-1 The Central Potential 132
 8-2 The Hydrogen Atom 133
 8-3 The Energy Spectrum 133
 The Degeneracy of the Spectrum 135
 8-4 The Free Particle 140
 8-5 Particle in an Infinite Spherical Well 143
 Problems

SUPPLEMENTS ON THE WEB SITE:

- 8-A A Useful Theorem** W-35
8-B The Square Well, Continuum Solutions W-38
 The Plane Wave Spherical Harmonics W-41

9. Matrix Representation of Operators 146

- 9-1 Matrices in Quantum Mechanics 146
 9-2 Matrix Representations of Angular Momentum Operators 150
 9-3 General Relations in Matrix Mechanics 151
 Problems 156

10. Spin 158

- 10-1 Eigenstates of Spin 1/2 158
 10-2 The Intrinsic Magnetic Moment of Spin 1/2 Particles 162
 10-3 Paramagnetic Resonance 164
 10-4 Addition of Two Spins 166
 10-5 The Addition of Spin 1/2 and Orbital Angular Momentum 171
 General Rules for Addition of Angular Momenta 171
 Problems 172

SUPPLEMENTS ON THE WEB SITE:

- 10-A The Addition of Spin 1/2 and Orbital Angular Momentum (Details)** W-43
 *General Rules for Addition of Angular Momenta, and Implications
 for Identical Particles* W-44
10-B The Levi-Civita Symbol and Maxwell's Equations W-47

11. Time-Independent Perturbation Theory 174

- 11-1 Energy Shifts and Perturbed Eigenstates 174
 11-2 Degenerate Perturbation Theory 177
 11-3 The Stark Effect 180
 The Stark Effect for $n = 2$ States 182
 Some Final Comments 184
 Problems 185

12. The Real Hydrogen Atom 188

12-1	Relativistic Kinetic Energy Effects	188
12-2	Spin-Orbit Coupling	188
12-3	The Anomalous Zeeman Effect	191
12-4	Hyperfine Structure	194
12-5	Comments on Reduced Mass Effects	196
	Problems	197

13. Many Particle Systems 199

13-1	The Two-Particle System	199
13-2	Identical Particles	201
	The Exchange Operator	201
13-3	The Pauli Principle	202
	N Fermions in a Potential Well	203
	When Is Antisymmetrization Necessary?	204
13-4	The Exclusion Principle and the Two-Particle Problem	206
	A Digression on Parity	206
13-5	The Exclusion Principle and Noninteracting Particles	207
13-6	Applications	211
	Degeneracy Pressure and Bulk Modulus	211
	Astrophysical Application	211
	Problems	213

SUPPLEMENT ON THE WEB SITE:

13-A	Conservation of Total Momentum	W-49
-------------	---------------------------------------	-------------

14. About Atoms and Molecules 215

14-1	The Helium Atom without Electron-Electron Repulsion	215
	Effects of the Exclusion Principle	217
14-2	Effects of Electron-Electron Repulsion	218
14-3	Exclusion Principle and Exchange Interaction	220
	Autoionization	223
14-4	The Ritz Variational Principle	224
	The Ground-State Energy of Helium	225
14-5	The Atom with Z Electrons	226
14-6	The Simplest Molecule H_2^+	228
	Molecular Orbitals	229
14-7	Molecular Spectra	231
	The Role of the Pauli Principle in Molecular Spectra	232
	Problems	233

SUPPLEMENTS ON THE WEB SITE:

14-A	The Hartree Approximation	W-51
14-B	The Building-Up Principle	W-55
	Spectroscopic Description of Ground States	W-58
14-C	A Brief Discussion of Molecules	W-64
	The H_2 Molecule	W-64
	The Valence Bond Method	W-66
	The Importance of Unpaired Valence Electrons	W-67

15. Time-Dependent Perturbation Theory 236

-
- 15-1 Formalism 236
 - 15-2 Harmonic Time-Variation of the Potential 238
 - 15-3 Phase Space 241
 - Problems 244

SUPPLEMENTS ON THE WEB SITE:**15-A Lifetimes, Linewidths, and Resonances W-69****15-B The Interaction Picture W-73****16. The Interaction of Charged Particles with the Electromagnetic Field 246**

-
- 16-1 Classical Electrodynamics 246
 - 16-2 The Schrödinger Equation for an Electron in Interaction with Electromagnetic Field 249
 - 16-3 The Constant Magnetic Field 250
 - Classical Limit 252
 - 16-4 Landau Levels 253
 - 16-5 The Integral Quantum Hall Effect 255
 - 16-6 An Additional Comment on Gauge Invariance 257
 - Problems 258

SUPPLEMENTS ON THE WEB SITE:**16-A The Aharov-Bohm Effect W-75****16-B A Little about Bessel Functions W-80****17. Radiative Decays 260**

-
- 17-1 Transition Rate 260
 - 17-2 Calculation of the Matrix Element 262
 - 17-3 Angular Integration and Selection Rules 263
 - 17-4 The $2p \rightarrow 1s$ Transition 267
 - Problems 269

SUPPLEMENT ON THE WEB SITE:**17-A Spin and Intensity Rules W-81****18. Selected Topics on Radiation 270**

-
- 18-1 Lasers 270
 - Conditions for the Operation of a Laser 271
 - Optical Pumping 271
 - The Cavity 272
 - 18-2 The Cooling of Atoms 273
 - 18-3 The Two-Level System 275
 - 18-4 The Three-Level System 278
 - 18-5 Observation of Quantum Jumps 279
 - 18-6 The Mössbauer Effect 282

SUPPLEMENTS ON THE WEB SITE:**18-A Quantizing the Electromagnetic Field without Frills W-83****18-B Details of the Three-Level System W-86****Dark States W-89****Electromagnetically Induced Transparency W-89**

19. Collision Theory 287

19-1	Collision Cross Section	287
	Elastic and Inelastic Scattering	291
19-2	Scattering at Low Energies	294
	The Breit-Wigner Formula	297
	S-Wave Scattering for Square Well	298
	Connection between Scattering Amplitude and Binding Energy	299
	Spin-Dependent Scattering	300
19-3	The Born Approximation	302
19-4	Scattering of Identical Particles	305
	Scattering by Atoms on a Lattice	307
19-5	An Inelastic Process: The Photoelectric Effect	308
	Problems	312

20. Entanglement and Its Implications 316

20-1	The Two-Slit Experiment and Which-way Information	318
20-2	The Quantum Eraser	320
20-3	Entanglement and EPR	322
20-4	Quantum Measurements	325
	Decoherence and the Schrödinger Cat	327

SUPPLEMENTS ON THE WEB SITE:

20-A	<i>Which-way Measurements and the Quantum Eraser</i>	W-91
20-B	<i>The Creation of GHZ States</i>	W-95
20-C	<i>The Density Operator</i>	W-98
	<i>Pure State</i>	W-98
	<i>Mixed State</i>	W-100

Physical Constants 329

References 330

Index 333

Chapter 1

The Emergence of Quantum Physics

Quantum theory grew out of an interplay of ground-breaking experiments and radical theoretical proposals that were not based on accepted classical physics. Our objective in this chapter is to summarize briefly the critical events that took place between 1900 and 1925, when quantum mechanics was created. We shall be very brief in our coverage, since these matters are discussed in much more detail in textbooks on modern physics.¹

1-1 BLACKBODY RADIATION

The study of radiation emitted by a heated body in equilibrium, initiated by the German physicist Gustav Kirchhoff in 1859, led to the recognition that the intensity of radiation emitted by a totally absorbing surface (a *black body*)—in practice a small hole in a heated enclosure—was a universal function of the temperature T and the wavelength λ (or frequency ν) of the radiation under consideration. For a cavity, the energy emitted per unit area, per unit time, in the form of this *blackbody radiation* in the wavelength range λ to $\lambda + d\lambda$, $E(\lambda, T)d\lambda$, is related to the energy density of radiation $w(\lambda, T)d\lambda$ inside the cavity

$$w(\lambda, T) = \frac{4E(\lambda, T)}{c} \quad (1-1)$$

where c is the speed of light. Wilhelm Wien showed in 1894 that classical thermodynamics led to the requirement that $w(\lambda, T)$ had to have the form

$$w(\lambda, T) = \frac{f(\lambda T)}{\lambda^5} \quad (1-2)$$

a result borne out by measurements. (See Fig. 1-1.) If the function $f(x)$ has a maximum, then at that maximum $\lambda T = x_{\max}$. Experiments determined that

$$\lambda_{\max} = \frac{b}{T} \quad (1-3)$$

where $b = 2.898 \times 10^{-3}$ m·K. In terms of the frequency ν , with $\nu = c/\lambda$, the energy density inside the cavity is

$$u(\nu, T) = w(\lambda, T) \left| \frac{d\lambda}{d\nu} \right| = w\left(\frac{c}{\nu}, T\right) \frac{c}{\nu^2} \quad (1-4)$$

¹See, for example, *Modern Physics* by J. Bernstein, P. M. Fishbane, and S. Gasiorowicz, Prentice Hall, Englewood Cliffs, N.J., 2000.

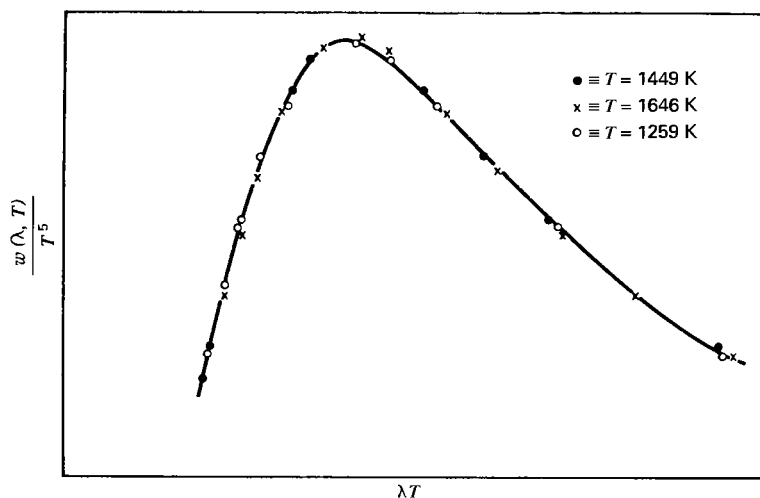


Figure 1-1 Experimental verification of (1-2) in the form of $w(\lambda, T)/T^5$ = a universal function of λT .

Painstaking experiments by O. Lummer and E. Pringsheim and by H. Rubens and F. Kurlbaum confirmed this result. Their data gave $u(\nu, T)$ at low and high frequencies (Fig. 1-2). Classical physics could not explain the shape of $f(\lambda, T)$. J. W. S. Rayleigh predicted a form of $u(\nu, T)$ based on equipartition of energy. If we assume that the average energy per degree of freedom for an oscillator is kT , and given that the number of modes (i.e., degrees of freedom) for standing waves of electromagnetic radiation in the cavity in the frequency range ν to $\nu + d\nu$ is $2 \times 4\pi\nu^2 d\nu / c^3$, the predicted result is²

$$u(\nu, T) = \frac{8\pi\nu^2}{c^3} kT \quad (1-5)$$

This result cannot, of course, be true for all frequencies, since the total energy $U(T)$ per unit volume obtained by integrating over all frequencies would be infinite!

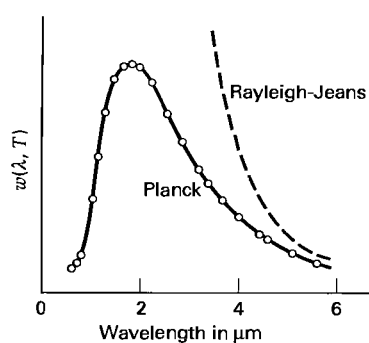


Figure 1-2 Comparison of the Planck formula and the Rayleigh-Jeans formula with data.

²A result of the form (1-5) is the only possible one, assuming that the size of the cavity does not enter into the expression and that the details of the matter that makes up the walls of the cavity are not important, so that e^2 and the electron (or ion) mass do not appear in the energy per unit frequency interval per unit volume. The only quantity with the dimensions of energy is kT , and the only length is $\lambda = c/\nu$.

At the end of 1900, the German physicist Max Planck obtained a formula for $u(\nu, T)$, which read

$$u(\nu, T) = \frac{8\pi h}{c^3} \frac{\nu^3}{e^{h\nu/kT} - 1} \quad (1-6)$$

where T is the temperature, k is the Boltzmann constant whose value is $k = 1.3807 \times 10^{-23}$ J/K, and h is a new constant whose value was obtained by fitting the data. Its value is $h = 6.6261 \times 10^{-34}$ J·s.

Planck obtained this formula by focusing on the dynamic equilibrium between the emission and absorption of radiation by the walls of the cavity, assumed to consist of simple oscillators. Only by making the radical assumption that *radiation of a given frequency ν could only be emitted and absorbed in “quanta” of energy given by $E = h\nu$* could he obtain the distribution in (1-6). A proper derivation involves material that we do not deal with at this stage.³ We present a very nice derivation of this formula in Supplement 1-A. [www.wiley.com/college/gasiorowicz] This derivation is due to Albert Einstein (1917), who used some of the ideas from the Bohr model, which we will discuss later in this chapter. At this point we note the following:

1. The total radiation energy per unit volume in the cavity is

$$U(T) = \frac{8\pi h}{c^3} \int_0^\infty d\nu \frac{\nu^3}{e^{h\nu/kT} - 1} = \frac{8\pi h}{c^3} \left(\frac{kT}{h}\right)^4 \int_0^\infty dx \frac{x^3}{e^x - 1} \quad (1-7)$$

The integral can be evaluated

$$\begin{aligned} \int_0^\infty dx \frac{x^3}{e^x - 1} &= \int_0^\infty dx x^3 e^{-x} \sum_{n=0}^\infty e^{-nx} = \sum_{n=0}^\infty \frac{1}{(n+1)^4} \int_0^\infty dy y^3 e^{-y} \\ &= 6 \sum_{n=1}^\infty \frac{1}{n^4} = \frac{\pi^4}{15} \end{aligned}$$

and the result is the Stefan-Boltzmann expression

$$U(T) = aT^4 \quad (1-8)$$

with $a = 7.5662 \times 10^{-16}$ J/m³·K⁴. The T^4 dependence had been derived earlier by Boltzmann, using general thermodynamic arguments, but the numerical value, depending as it does on the value of h , was beyond the reach of classical physics.

2. The value of b in eq. (1-3) can be calculated and agrees with experiment.
3. It is only at low frequencies that $u(\nu, T)$ takes the form predicted by Rayleigh (who made his prediction for *all* frequencies!) in eq. (1-5), in agreement with experiment.
4. At high frequencies, the rise in the number of modes is more than compensated for by the drop in the average energy modulated by the exponentially falling factor $e^{-h\nu/kT}$. In effect, the Planck formula showed that the *quantization* implied that the energy per degree of freedom was not the frequency-independent classical equipartition value kT , but rather an energy that was much smaller at high frequencies.

³This is discussed in Bernstein et al., *loc. cit.*

EXAMPLE 1-1

Assume that the sun radiates as a black body. Given that the radius of the sun is 7×10^8 m, (a) what is the total amount of radiation emitted per second? (b) Given that the distance between the earth and the sun is 1.5×10^{11} m, how much energy falls on 1 m^2 at earth on a surface perpendicular to the direction to the sun? You may use eqs. (1-1) and (1-7) and assume that the temperature of the surface of the sun is 6000 K.

SOLUTION

(a) Just as the energy density is related to the emissive power by (1-1), the amount of energy radiated by a surface is given by

$$\begin{aligned} E(T) &= \frac{c}{4} U(T) = \frac{1}{4} (3 \times 10^8 \text{ m/s}) (7.57 \times 10^{-16} \text{ J/m}^3 \cdot \text{K}^4) (6000 \text{ K})^4 \\ &= 7.36 \times 10^7 \text{ W/m}^2 \end{aligned}$$

This is to be multiplied by the surface area of the sun

$$A = 4\pi R^2 = 4\pi (7 \times 10^8 \text{ m})^2 = 6.2 \times 10^{18} \text{ m}^2$$

This gives us the total power radiated

$$P = E(T)A = 4.5 \times 10^{25} \text{ W}$$

(b) The total radiation falls on the surface of a sphere of radius 1.5×10^{11} m, so that a fraction $1/(4\pi D^2) = 1/(4\pi \times (1.5 \times 10^{11} \text{ m})^2)$ is intercepted by a single square meter. This gives

$$p = \frac{P}{4\pi D^2} = \frac{P}{4\pi (1.5 \times 10^{11} \text{ m})^2} = 1.6 \times 10^3 \text{ W}$$

Note: The observed number, the *solar constant*, is $1.36 \times 10^3 \text{ W}$.

EXAMPLE 1-2

Use dimensional analysis to get an expression for the parameter a in the Stefan-Boltzmann law.

SOLUTION The relation that defines a is $U(T) = aT^4$. Actually, temperature is dimensionless, but kT is an energy. Thus we really should write this as $U(T) = \alpha(kT)^4$. We see that α has the dimensions of (energy)/(length)³ (energy)⁴ = (1/energy $\times$ length)³. The only quantities that can enter here are h and c . The electron mass, its charge, and so on are irrelevant here. Now $(h)^p(c)^q = (ML^2/t)^p(L/t)^q = (1/ML^3/t^2)^3$. We immediately see that $p = -3$, to match the M dimensions. We then get $L^{-6+q}t^{2-2q} = L^{-9}t^6$. We match these with $q = -3$. Hence we can write $\alpha = (\text{const})(hc)^{-3}$, and therefore $a = (\text{const}) \frac{k^4}{(hc)^3}$.

EXAMPLE 1-3

Use dimensional analysis to get an expression for the parameter b in the Wien relation $b = \lambda_{\max}T$.

SOLUTION We see that $b = (1/k)(\lambda_{\max})(kT)$. Now the second and third terms together have the dimensions of (length $\times$ energy), and we saw in Example 1-2 that this can be constructed out of (hc) . Again, the electron charge and mass do not enter; we therefore get $b = (\text{const}) \frac{hc}{k}$.

In 1905, Albert Einstein published an extremely important paper in which he departed from the Planckian notion that it was the absorption and emission of radiation that occurred in quanta. Einstein instead proposed that *radiation itself consisted of quanta of energy*, with the rule that the energy of a quantum of light of frequency ν is $h\nu$. Einstein then used the concept of the quantum nature of light to explain some peculiar properties of metals when these are irradiated with visible and ultraviolet light.

1-2 THE PHOTOELECTRIC EFFECT

In 1887, the photoelectric effect was discovered by Heinrich Hertz, who, while engaged in his famous experiments on electromagnetic waves, found that the length of the spark induced in the secondary circuit was reduced when the terminals of the spark gap were shielded from the ultraviolet light coming from the spark in the primary circuit. His observations attracted much interest, and the following facts were established by further experiments:

1. When polished metal plates are irradiated, they may emit electrons;⁴ they do not emit positive ions.
2. Whether the plates emit electrons depends on the wavelength of the light. In general there will be a threshold that varies from metal to metal: Only light with a frequency greater than a given threshold frequency will produce a photoelectric current.
3. The magnitude of the current, when it exists, is proportional to the intensity of the light source.
4. The energy of the photoelectrons is independent of the intensity of the light source but varies linearly with the frequency of the incident light.

Although the existence of the photoelectric effect can be understood within the framework of classical electromagnetic theory, since it was known that there were electrons in metals, and one could imagine them to be accelerated by absorption of radiation, the frequency-dependence of the effect is not comprehensible within that framework. The energy carried by an electromagnetic wave is proportional to the intensity of the source, and frequency has nothing to do with it. Furthermore, a classical explanation of the effect, which would have to involve the concentration of the energy deposited on single photoelectrons, would carry with it an implied time delay between the arrival of the radiation and the departure of the electron, the delay being longer when the intensity is decreased. In fact, no such time delays were ever observed, at least none longer than 10^{-9} sec, even with incident radiation of very low intensity.

Einstein considered the implications of having the radiation consist of a collection of quanta of energy $h\nu$, where ν is the frequency of the light wave. The absorption of a single quantum of radiation by an electron—a process that may take less time than the upper limit previously quoted—increases the electron energy by an amount $h\nu$. Some of the energy must be expended to separate the electron from the metal. This amount, W (called the *work function*), might be expected to vary from metal to metal, but should not depend on the electron energy. The rest is available for the electron kinetic energy, so that on the

⁴This was established by an e/m measurement.

basis of this picture one expects that the following relation between electron velocity v and light frequency ν .

$$\frac{1}{2}mv^2 = h\nu - W \quad (1-9)$$

should hold. The threshold effect and the linear relation between electron kinetic energy and the frequency are contained in this formula. The proportionality of the current and the source intensity can also be understood in terms of these light quanta, or *photons*, as they came to be called: a more intense light source emits more photons, and these in turn can liberate more electrons.

The confirmation of Einstein's hypothesis was slow in coming. The definitive experiments carried out by the American physicist Robert A. Millikan (who did not believe in quanta!) were not completed until 1915. As Fig. 1-3 shows, they were in complete accord with (1-9). Incidentally, a byproduct of these and subsequent experiments was information about metals. It was found that W was of the order of several electron volts (1 eV is the energy required to move a charge e through one volt potential difference, and since the charge of the electron is $|e| = 1.602 \times 10^{-19}$ C, we have 1 eV = 1.602×10^{-19} J).

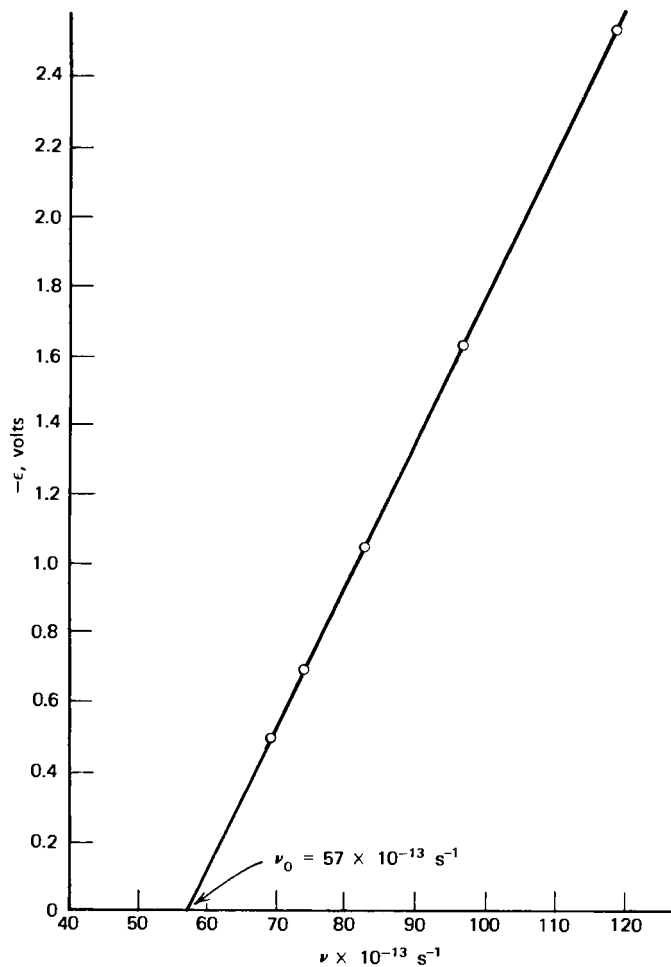


Figure 1-3 Photoelectric effect data showing a plot of retarding potential necessary to stop electron flow from a metal (lithium), or equivalently, electron kinetic energy, as a function of frequency of the incident light. The slope of the line is h/e .

EXAMPLE 1-4

Radiation of wavelength $\lambda = 290 \text{ nm}$ falls on a metal surface for which the work function is $W = 4.05 \text{ eV}$. What potential is needed to stop the most energetic photoelectrons?

SOLUTION The energy of the photons is

$$\begin{aligned} E = h\nu = hc/\lambda &= (6.63 \times 10^{-34} \text{ J}\cdot\text{s})(3 \times 10^8 \text{ m/s})/(2.9 \times 10^{-7} \text{ m}) = 6.86 \times 10^{-19} \text{ J} \\ &= (6.86 \times 10^{-19} \text{ J})/(1.60 \times 10^{-19} \text{ J/eV}) = 4.28 \text{ eV} \end{aligned}$$

The maximum kinetic energy of the photoelectron is $E - W = 0.23 \text{ eV}$, so the potential difference is 0.23 volts.

1-3 THE COMPTON EFFECT

Einstein's proposal that the quanta of energy should also manifest themselves outside of the "blackbody cavity," and the attendant explanation of the photoelectric effect, did not yet lead him to the notion of the quantum of radiation as a *particle*. It was not until 1916 that Einstein further analyzed the concept of a *gas of quanta* and was led to the prediction that the quanta, hereafter called **photons**, had to carry momentum as well as energy. The experiments that provided the most direct confirmation of the particle nature of radiation were carried out by Arthur H. Compton in 1923–4. Compton discovered that radiation of a given wavelength (in the X-ray region) sent through a metallic foil was scattered in a manner not consistent with classical radiation theory. According to classical theory, the mechanism for the effect is the re-radiation of light by electrons set into forced oscillations by the incident radiation, and this leads to the prediction of intensity observed at an angle θ that varies as $(1 + \cos^2 \theta)$, and does not depend on the wavelength of the incident radiation. Compton found that the radiation scattered through a given angle actually consists of two components: one whose wavelength is the same as that of the incident radiation, the other of wavelength shifted relative to the incident wavelength by an amount that depends on the angle (Fig. 1-4). Compton was able to explain the "modified" component by treating the incoming radiation as a beam of photons of energy $h\nu$, with individual photons scattering elastically off individual electrons. In an elastic collision, momentum

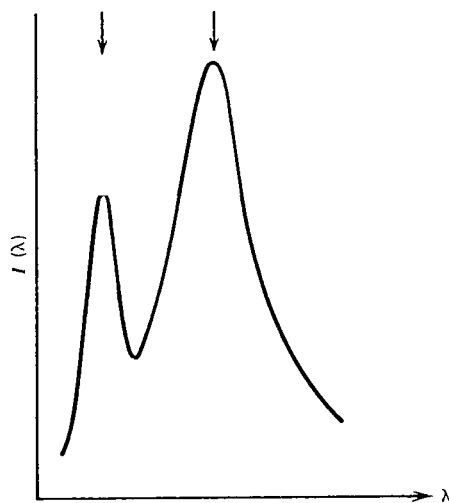


Figure 1-4 The spectrum of radiation scattered by carbon, showing the unmodified line at 0.07078 nm on the left and the shifted line at 0.07314 nm on the right. The former is the wavelength of the primary radiation.

as well as energy must be conserved, and we must first assign a momentum to the photon. By analogy with relativistic particle kinematics we argue that

$$p = \frac{h\nu}{c} \quad (1-10)$$

The argument is that it follows from the relativistic relation between energy and momentum

$$E = [(m_0c^2)^2 + (pc)^2]^{1/2} \quad (1-11)$$

where m_0 is the rest mass of the particle, that the velocity at this momentum is

$$v = \frac{dE}{dp} = \frac{pc^2}{E} = \frac{pc^2}{(m_0^2c^4 + p^2c^2)^{1/2}} \quad (1-12)$$

For a photon this is always c , and hence the *photon rest mass must be zero*. Thus the relation (1-11) becomes

$$E = pc \quad (1-13)$$

which yields (1-10) when we substitute $E = h\nu$. One may also derive (1-13) from consideration of the energy and momentum of an electromagnetic wave, but the analogy argument is simpler.

Consider, now, a photon with initial momentum $\mathbf{p}$ incident upon an electron at rest. After the collision, the photon momentum is $\mathbf{p}'$, and the electron recoils with momentum $\mathbf{P}$. Conservation of momentum yields (Fig. 1-5)

$$\mathbf{p} = \mathbf{p}' + \mathbf{P} \quad (1-14)$$

from which it follows that

$$\mathbf{P}^2 = (\mathbf{p} - \mathbf{p}')^2 = \mathbf{p}^2 + \mathbf{p}'^2 - 2\mathbf{p} \cdot \mathbf{p}' \quad (1-15)$$

Energy conservation reads

$$h\nu + m_e c^2 = h\nu' + (m_e^2 c^4 + P^2 c^2)^{1/2} \quad (1-16)$$

where m_e is the electron mass. When this is rearranged and squared, one finds that

$$\begin{aligned} m_e^2 c^4 + P^2 c^2 &= (h\nu - h\nu' + m_e c^2)^2 \\ &= (h\nu - h\nu')^2 + 2m_e c^2(h\nu - h\nu') + m_e^2 c^4 \end{aligned}$$

On the other hand, (1-15) may be rewritten in the form

$$P^2 = \left(\frac{h\nu}{c}\right)^2 + \left(\frac{h\nu'}{c}\right)^2 - 2\frac{h\nu}{c} \cdot \frac{h\nu'}{c} \cos \theta$$

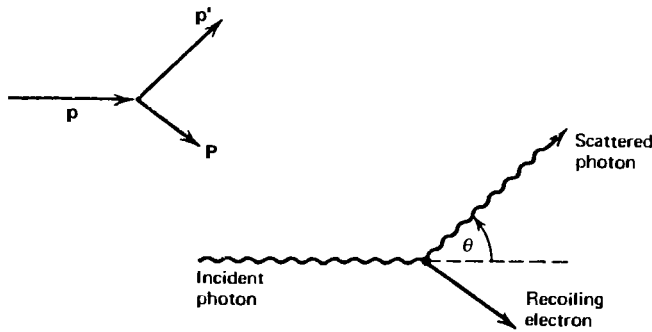


Figure 1-5 Kinematics for the Compton effect.

that is,

$$P^2c^2 = (h\nu - h\nu')^2 + 2(h\nu)(h\nu')(1 - \cos \theta) \quad (1-17)$$

where θ is the photon scattering angle. If we now insert P^2c^2 obtained from the energy conservation equation, we are led to

$$h\nu\nu'(1 - \cos \theta) = m_e c^2(\nu - \nu')$$

or equivalently

$$\lambda' - \lambda = \frac{h}{m_e c} (1 - \cos \theta) \quad (1-18)$$

The measurements of the modified component agree very well with this prediction. The unmodified line is due to the scattering by the whole atom; if m_e is replaced by the mass of the atom, the shift in the wavelength is very small, since an atom is many thousand times more massive than an electron. The quantity $h/m_e c$ has the dimensions of a length. It is called the *Compton wavelength* of the electron, and its magnitude is

$$\frac{h}{m_e c} = 2.426 \times 10^{-12} \text{ m} \quad (1-19)$$

Measurements of the electron recoil were also made, and these are in good agreement with the kinematic calculation outlined above. It was furthermore determined by good time resolution coincidence experiments that the outgoing photon and the recoil electron appear simultaneously. Thus there is no question of the correctness of the interpretation of the Compton effect as an elastic collision between a massive and a massless object. The particle-like behavior of the photon is not obvious to the naked eye. Nevertheless, it is interesting to note that although the energy of a quantum of radiation in the optical range, with $\lambda = 450 \text{ nm}$, has a very tiny energy,

$$h\nu = hc/\lambda = (6.63 \times 10^{-34} \text{ J}\cdot\text{s})(3 \times 10^8 \text{ m/s})/(4.5 \times 10^{-7} \text{ m}) = 4.24 \times 10^{-19} \text{ J},$$

the dark-adapted eye can detect as few as 10 photons. A light bulb emitting 20 J/s of energy in that wavelength vicinity will emit

$$n = (20 \text{ J/s})/(4.24 \times 10^{-19} \text{ J}) = 4.7 \times 10^{19} \text{ photons/second}$$

The real difficulty arises from the knowledge that radiation also has wave properties and exhibits interference and diffraction. This gives rise to enormous conceptual difficulties, which will be addressed in Chapter 2; the resolution of these difficulties finally came with quantum mechanics.

EXAMPLE 1-5

A photon of energy 500 MeV strikes a proton at rest. What is its maximum energy loss?

SOLUTION The maximum momentum transfer to the photon occurs when the proton moves in the direction of the incident photon—a head-on collision. We can thus deal with this problem as a one-dimensional one. Let the initial photon momentum be $+q$, and let its final momentum be q' . Since the proton was initially at rest, its final momentum must be $q - q'$. Energy conservation gives us

$$qc + Mc^2 = |q'|c + \sqrt{(Mc^2)^2 + (q - q')^2 c^2}$$

The absolute sign comes in because we don't know the sign of q' . Suppose $q' > 0$. Then

$$(qc - q'c + Mc^2)^2 = (Mc^2)^2 + (q - q')^2 c^2$$

which implies that $2Mc^3(q - q') = 0$; that is, $q' = q$. The proton remains at rest and there is no energy loss.

Suppose $q' < 0$. In that case the final proton momentum is $q + |q'|$. Energy conservation now implies that

$$((q - |q'|)c + Mc^2)^2 = (Mc^2)^2 + (q + |q'|)^2 c^2$$

After division by c^2 this leads to

$$2(q - |q'|)Mc = q^2 + 2q|q'| + q'^2 - q^2 + 2q|q'| - q'^2 = 4q|q'|$$

Hence

$$|q'| = \frac{q}{1 + 2q/Mc}$$

The energy loss of the photon is

$$\Delta E = (q - |q'|)c = \frac{2q^2}{M} \frac{1}{1 + 2q/Mc}$$

With $qc = 500$ MeV and $Mc^2 = 936$ MeV, we get $\Delta E = 260$ MeV.

1-4 WAVE PROPERTIES AND ELECTRON DIFFRACTION

In 1923 de Broglie, guided by the analogy of Fermat's principle in optics and the least-action principle in mechanics, was led to suggest that the dual wave-particle nature of radiation should have its counterpart in a dual particle-wave nature of matter. Thus particles should have wave properties under certain circumstances, and de Broglie suggested an expression for the wavelength associated with the particle.⁵ This is given by

$$\lambda = \frac{h}{p} \quad (1-20)$$

where h is Planck's constant and p is the momentum of the particle. De Broglie's work attracted much attention, and many people suggested that verification could be obtained by observing electron diffraction.⁶ The experimental observation of electron diffraction by a regular array of atoms forming the lattice of a crystal was carried out by C. J. Davisson and L. H. Germer and independently by G. P. Thomson, who found that there was preferential scattering in certain directions.

Figure 1-6 is a simplified picture of what happens. When waves are scattered by a periodic structure, there will be a phase difference between waves coming from adjacent scattering "planes." The phase difference is given by $(2\pi/\lambda)2a \sin \theta$, and there will be constructive interference whenever the phase difference equal $2\pi n$, where n is an integer. The result

$$n\lambda = 2a \sin \theta \quad (1-21)$$

⁵The relation parallels the photon relation $\lambda = c/\nu = hc/h\nu = hc/E = h/p$.

⁶The history of the verification of de Broglie's conjecture can be found in Max Jammer, *The Conceptual Development of Quantum Mechanics*. Los Angeles: Tomash Publishers; New York: American Institute of Physics, 1989 (originally published 1966).

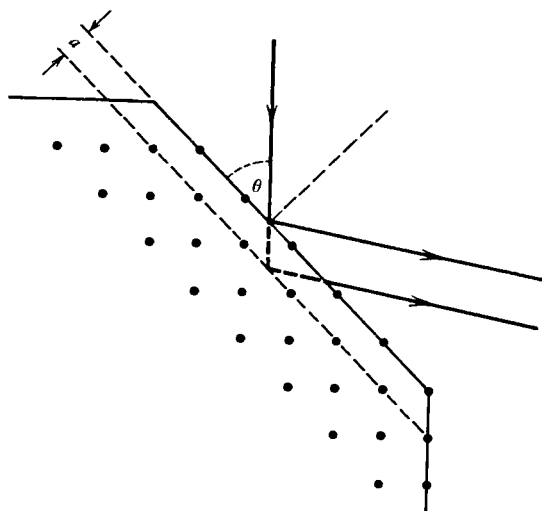


Figure 1-6 Schematic drawing of electron scattering geometry.

is called the Bragg condition, because it was discovered by W. L. Bragg in 1912 in connection with his work on the scattering of X rays by crystals. The interference pattern observed in electron scattering could be correlated with (1-21) provided the association (1-20) between the energy of the electrons and the wavelength was made. This verification constituted a major development in wave mechanics, and led Schrödinger to his equation.

EXAMPLE 1-6

What is the de Broglie wavelength of an electron whose energy is 6.0 eV? What is it for an electron of energy 200 MeV? (The mass of an electron is $0.51 \text{ MeV}/c^2$.)

SOLUTION In the first case we have a nonrelativistic electron. Hence $E = p^2/2m_e = 6.0 \text{ eV}$. This allows us to write

$$(pc)^2 = 2m_e c^2 (6.0 \text{ eV}) = 2(0.51 \times 10^6 \text{ eV})(6 \text{ eV})$$

so that

$$pc = 2.47 \times 10^3 \text{ eV}$$

Hence $\lambda = h/p = hc/pc$. Now

$$\begin{aligned} hc &= (6.63 \times 10^{-34} \text{ J}\cdot\text{s})(3 \times 10^8 \text{ m/s}) = 1.99 \times 10^{-25} \text{ J}\cdot\text{m} \\ &= (1.99 \times 10^{-25} \text{ J}\cdot\text{m})/(1.60 \times 10^{-19} \text{ J/eV}) = 1.24 \times 10^{-6} \text{ eV}\cdot\text{m} \end{aligned}$$

Thus in the first case, $\lambda = (1.24 \times 10^{-6} \text{ eV}\cdot\text{m})/(2.47 \times 10^3 \text{ eV}) = 0.50 \times 10^{-9} \text{ m}$. In the second case, the energy is highly relativistic, so that $E = pc = 200 \times 10^6 \text{ eV}$. Thus $\lambda = hc/pc = (1.24 \times 10^{-6} \text{ eV}\cdot\text{m})/(2 \times 10^8 \text{ eV}) = 0.62 \times 10^{-14} \text{ m}$. This is comparable to the size of a nucleus.

A wonderfully instructive experiment on electron diffraction by double slit was carried out by A. Tonomura. Figure 1-7 shows the experimental set-up. The source emits electrons one by one, and these pass between two parallel plates, with a thin filament ($1 \mu\text{m}$ in diameter) in between. This device is called a biprism. When the filament is

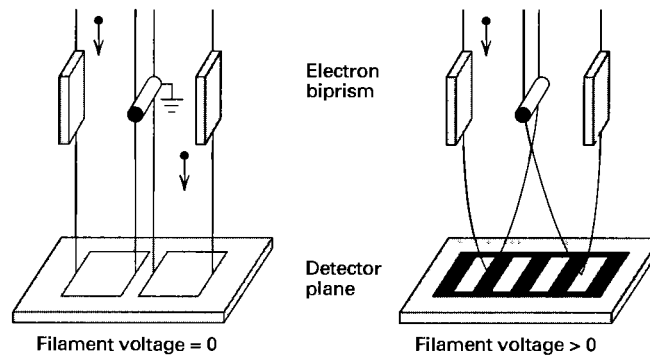


Figure 1-7 (a) When the filament is grounded, the electrons passing one by one can be seen on the monitor, and the path can be identified. The shadow of the filament can be seen after many electrons have passed. (b) When a positive voltage is applied to the filament, the electrons are bent toward the center, and there is no way of telling which side of the filament the electron passed by. An interference pattern results.

(Courtesy of A. Tonomura, Hitachi Advanced Research Laboratory)

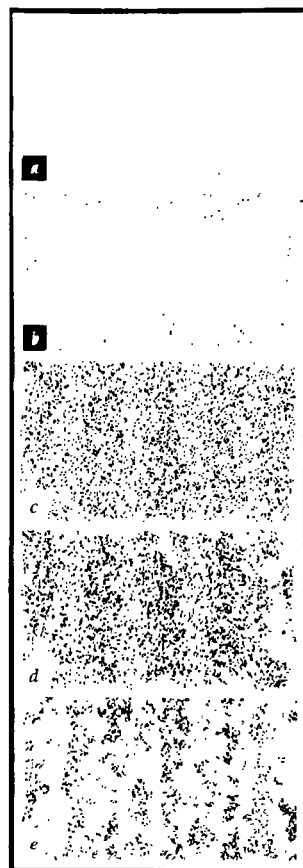


Figure 1-8 Patterns of dots arriving at the monitoring screen with (a) 10, (b) 100, (c) 3000, (d) 20,000, and (e) 70,000 electrons accumulating on the screen.

(Courtesy of A. Tonomura, Hitachi Advanced Research Laboratory)

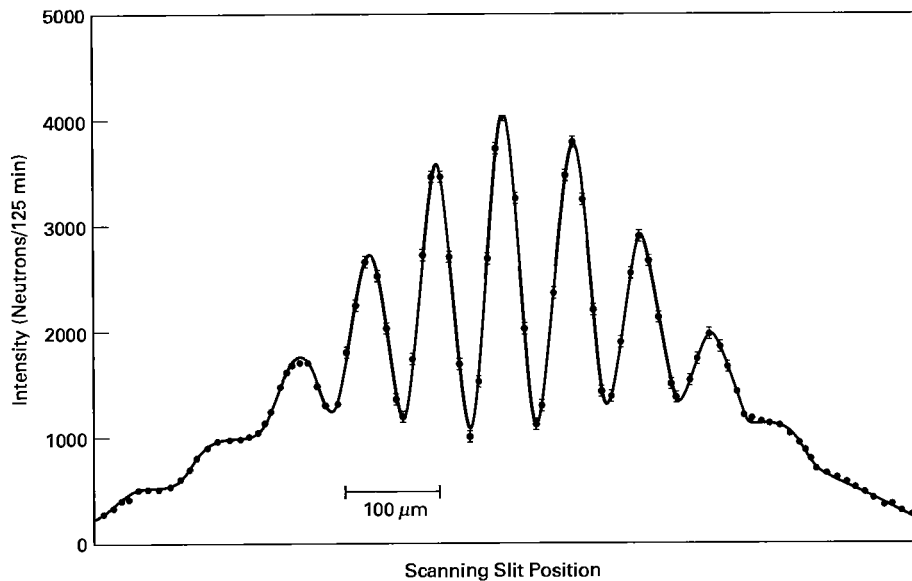


Figure 1-9 Double slit diffraction pattern for neutrons with wavelength $\lambda \approx 18.5 \text{ \AA}$.

(From A. Zeilinger, R. Gahler, C. G. Shull, W. Treimer, and W. Mampe, *Rev. Mod. Phys.* **60**:1067 (1988), by permission.)

grounded, the electrons pass on both sides of the filament and hit the detector. When a large number of electrons have arrived, the shadow of the filament is apparent in a uniform distribution of “hits.” When a positive voltage is applied to the filament, the electrons on the two sides are deflected by the filament, and the biprism forms the equivalent of a double slit. Figure 1-8 shows the pattern of hits on the detector. For a large number of electrons the interference pattern becomes evident. Furthermore, when the path of electrons on one side of the filament is blocked, the distribution of electrons passing on the other side is again uniform. We will discuss the implications of this experiment in some detail in Chapter 2. Suffice it to say that the wave properties of electrons are manifest in Fig. 1-8 (e).

Particle diffraction experiments have been carried out with slow neutrons (see Fig. 1-9), and with molecular beams of hydrogen and helium. At the time of writing, the most massive particles involved in a diffraction experiment have been *fullerenes*, consisting of 60 carbon atoms.⁷ Perhaps the most intriguing exhibition of particle-wave duality can be seen in the *Kapitza-Dirac* effect. In 1933 P. Kapitza and P. A. M. Dirac noted that electrons could be diffracted by the periodic structure set up by a standing electromagnetic wave. The experiment did not become feasible until the invention of the laser. The most effective experiment that clearly shows the expected diffraction peaks was carried out in 2001. Figure 1-10 shows the apparatus and Fig. 1-11 shows the results of the experiment.

Neutron diffraction has become an important tool in the study of crystal structure. Since the typical dimensions are of the order of Ångströms ($1 \text{ \AA} = 0.1 \text{ nm}$), the momen-

⁷The artistry involved in carrying out what is conceptually a simple experiment is described in M. Arndt et al., *Nature*, vol. 401 (1999), p. 680.

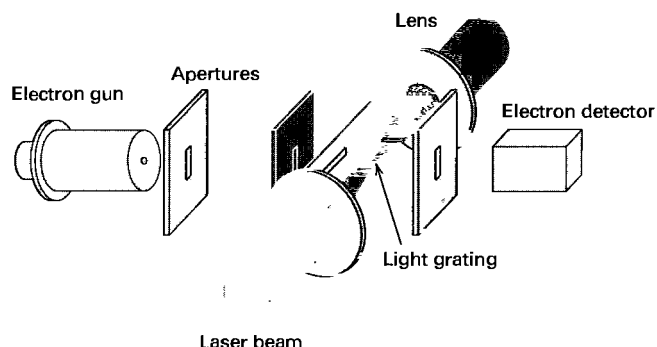


Figure 1-10 An electron beam is collimated and diffracts from a standing wave of radiation, which acts as a grating.

(Courtesy of D. L. Freimund, K. Aflatooni and H. Batelaan. Reprinted by permission from *Nature*, **413**, 142–143 (2001), copyright 2001, MacMillan Publishers Ltd.)

tum of a neutron corresponding to a de Broglie wavelength of 0.1 nm is $P = h/\lambda = 6.6 \times 10^{24} \text{ kg}\cdot\text{m/s}$, and for neutrons this leads to a kinetic energy of $1.3 \times 10^{-24} \text{ J}$, or equivalently 0.08 eV.

On a macroscopic scale it is beyond our ability to observe these wave aspects. A droplet of water 0.1 mm in diameter, moving at 1 mm/s, is easily seen to have a de Broglie wavelength of the order of 10^{-25} m , which is ten orders of magnitude smaller than the size of the proton. It is the smallness of Planck's constant that makes the separation of wave and particle properties so striking in the classical domain.

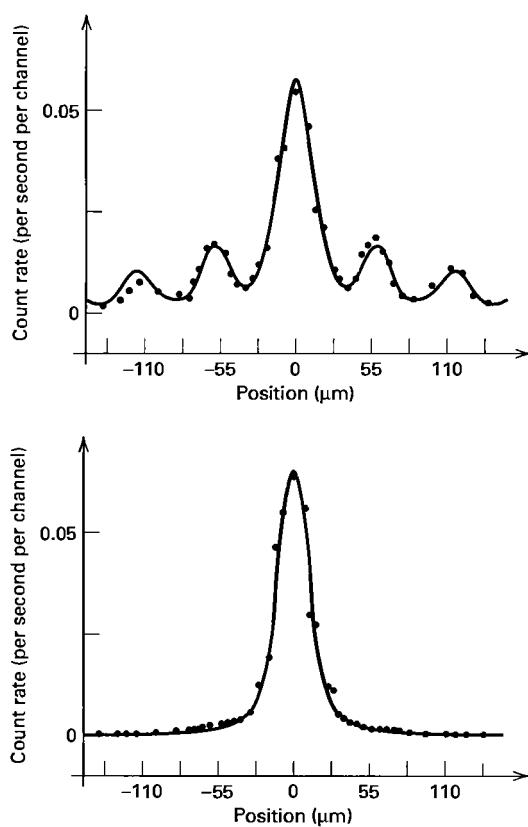


Figure 1-11 Electron detection rate with laser on and off. The solid curve is a result of a calculation using the Schrödinger equation.

(Courtesy of D. L. Freimund, K. Aflatooni and H. Batelaan. Reprinted by permission from *Nature*, **413**, 142–143 (2001), copyright 2001, MacMillan Publishers Ltd.)

1-5 THE BOHR ATOM

The Rutherford Planetary Model

The discovery of radioactivity by Henri Becquerel in 1896 provided the tools for an attack on the structure of the atom, which was complementary to the study of the emission of radiation by atoms. Ernest Rutherford was the leading physicist in the study of atomic structure, and he pioneered the use of the particles emitted in radioactive decay as projectiles to study atoms. Experiments carried out in 1908 under Rutherford's guidance by H. W. Geiger and E. Marsden, in which α -particles were aimed at thin foils, showed that a surprisingly large fraction of the α -particles underwent large-angle scattering, a result totally inconsistent with expectations based on the Thomson model of the atom. In the Thomson model, electrons were assumed to be embedded in a distribution of positive charge whose extent determined the atomic radius. Electrons do not deflect α -particles, since they are about 10^4 times less massive. Thus the source of deflection of the α -particles has to be the positive charge, and large-angle deflection implies that the potential at the surface of the charge distribution is large. This in turn implies that the positive charge is limited to a region much smaller than the volume of the atom. Rutherford proposed a new model that accounted for the data. In this model, all of the positive charge (and almost all of the mass) is concentrated in a small region in the center of the atom. This positively charged *nucleus* attracts the negatively charged electrons, and since the force law has a $1/r^2$ behavior, the *electrons travel in circular or elliptic orbits about the nucleus*.

The Rutherford model explained the α -scattering data. The *planetary orbit* picture of the electrons in atoms was an immediate consequence of the model, and it immediately faced insuperable difficulties. It is a consequence of classical electrodynamics that an accelerating charge radiates. Thus an electron in an orbit will continually lose energy, and one can estimate that it would take an electron approximately 10^{-10} s to spiral into the nucleus from an orbit of the size of 10^{-10} m. (See Supplement 1-B: www.wiley.com/college/gasiorowicz.) Furthermore, classical electrodynamics shows that the radiation emitted by charges in periodic motion has the frequency of that motion. This was in complete contradiction with a vast amount of information about atomic spectra known at the time.⁸ The information about the spectra of hydrogen, obtained by Anders Ångström and others, was summarized by the Swiss Johann Balmer in the formula discovered in 1885, whose modern form is

$$\frac{1}{\lambda} = Ry \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad (1-22)$$

where n_1 and n_2 are integers, and Ry is a constant.

The Bohr Postulates

Just two years after this model was proposed, the Danish physicist Niels Bohr in 1913 advanced a series of postulates, which, while sharply breaking with classical theory, explained the spectral structure and bypassed the stability problem. Bohr proposed that

1. An atomic system can only exist in a discrete set of *stationary states*, with discrete values of energy, and any change of the energy of the system, including emission and absorption of electromagnetic radiation must take place by a complete transition between two such stationary states.

⁸The reader will find a fascinating account in Chapter 9 of *Inward Bound* by Abraham Pais, Oxford University Press, England, 1986.

2. The radiation absorbed or emitted during a transition between two stationary states of energies E_1 and E_2 (with $E_1 > E_2$) is characterized by a unique frequency given by

$$h\nu = E_1 - E_2 \quad (1-23)$$

3. The stationary states correspond to a set of allowed orbits in the Rutherford atomic model. They are determined by the requirement that the kinetic energy of the electron in the orbit is related to the frequency f of the motion of the electron in that orbit by

$$\frac{1}{2} m_e v^2 = \frac{1}{2} nhf \quad (1-24)$$

where $n = 1, 2, 3, \dots$. For circular orbits⁹ this reduces to the statement that the *angular momentum takes on integer values in units of $(h/2\pi)$* , so that

$$mvr = n \frac{h}{2\pi} \quad (1-25)$$

The consequences of these postulates are very simply deduced for one-electron atoms such as hydrogen, singly ionized helium, and so on, if we deal with the circular orbits. If the nuclear charge is Ze and that of the electron is $-e$, and if the radius of the orbit is r , then, taking the nuclear mass to be infinite, we balance the Coulomb force against the centrifugal force

$$\frac{Ze^2}{4\pi\epsilon_0 r^2} = \frac{m_e v^2}{r} \quad (1-26)$$

This, when combined with (1-25), leads to

$$v = \frac{2\pi Ze^2}{4\pi\epsilon_0} \frac{1}{hn} \quad (1-27)$$

and

$$r = \frac{n^2 h^2}{4\pi^2} \frac{1}{(Ze^2/4\pi\epsilon_0)} \frac{1}{m_e} \quad (1-28)$$

The energy is

$$E = \frac{1}{2} m_e v^2 - \frac{Ze^2}{4\pi\epsilon_0 r} = -\frac{1}{2} \left(\frac{2\pi}{h} \right)^2 \left(\frac{Ze^2}{4\pi\epsilon_0} \right)^2 \frac{1}{n^2} \quad (1-29)$$

which, by postulate (2) immediately leads to the general form (1-22) (Fig. 1-12).

Before evaluating these quantities to obtain an idea of their magnitude, we will introduce some notations that will be very useful. First, it is $h/2\pi$ rather than h that appears in most formulas in quantum mechanics. We therefore define

$$\hbar = \frac{h}{2\pi} = 1.0546 \times 10^{-34} \text{ J}\cdot\text{s} \quad (1-30)$$

To keep the expressions for the energy simple, we shall deal with the angular frequency ω , rather than ν , where

$$\omega = 2\pi\nu \quad (1-31)$$

⁹When elliptical orbits are allowed, a much richer structure emerges. This will be discussed in Chapter 8, in the framework of the Schrödinger equation.

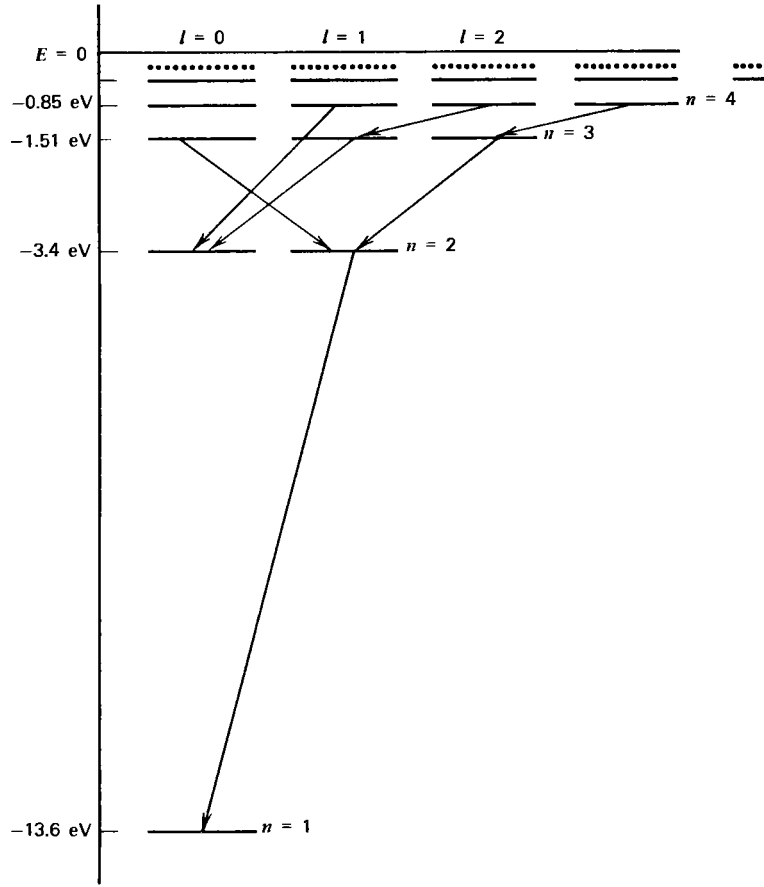


Figure 1-12 Spectrum for the hydrogen atom as derived from the Bohr atomic model. The existence of the quantum numbers l emerges from a discussion of elliptical orbits. The lines connecting energy levels represent some of the dominant atomic transitions.

Thus (1-23) reads

$$\omega = \frac{E - E'}{\hbar} \quad (1-32)$$

Similarly, the quantum of radiation carries energy

$$E = \hbar\omega \quad (1-33)$$

The Bohr angular momentum quantization condition reads

$$mvr = n\hbar \quad (n = 1, 2, 3, \dots) \quad (1-34)$$

It is also very convenient to introduce the dimensionless “fine structure constant”

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137.036} \quad (1-35)$$

which we will approximate by $1/137$. In terms of these quantities we find the much simpler expressions

$$v = \frac{Z\alpha c}{n} \quad r = \frac{\hbar}{m_e c Z \alpha} n^2 \quad (1-36)$$

and

$$E = -\frac{1}{2} m_e \frac{(Z\alpha c)^2}{n^2} \quad (1-37)$$

Notice that the radius, which has the dimensions of a length, is written in terms of $\hbar/mc$, the reduced Compton wavelength of the electron, and that the energy is written in terms of mc^2 . In all atomic calculations we shall express our results in terms of mc^2 , $\hbar/mc$, $\hbar/mc^2$, and mc for energy, length, time, and momentum, respectively. Angular momenta will always appear as multiples of $\hbar$.

Let us now calculate some of the quantities that emerge from the Bohr theory. We calculate

$$\begin{aligned} m_e c^2 &= 0.51 \times 10^6 \text{ eV} = 0.51 \text{ MeV} \\ \frac{\hbar}{m_e c} &= 3.9 \times 10^{-13} \text{ m} \\ \frac{\hbar}{m_e c^2} &= 1.3 \times 10^{-21} \text{ s} \end{aligned} \quad (1-38)$$

and thus obtain:

- (a) The radius of the lowest ($n = 1$) Bohr orbit is

$$a_0 = \frac{137}{Z} \frac{\hbar}{m_e c} = \frac{0.053}{Z} \text{ nm} \equiv \frac{0.53}{Z} \text{ \AA} \quad (1-39)$$

We have introduced here the useful unit of length, the Ångstrom, where $1 \text{ \AA} = 10^{-10} \text{ m} = 0.1 \text{ nm}$

- (b) The binding energy of the electron in the lowest Bohr orbit—that is, the energy required to put it in a state with $E = 0$ (corresponding to $n = \infty$)—is

$$E = \frac{1}{2} m_e c^2 (Z\alpha)^2 = 13.6 Z^2 \text{ eV} \quad (1-40)$$

Thus, for example, a transition from the $n = 1$ state to the $n = 2$ state in hydrogen ($Z = 1$) corresponds to a change in energy of $13.6 (1 - \frac{1}{4}) \text{ eV} = 10.2 \text{ eV}$. The frequency of the emitted radiation can be calculated by converting this into Joules, but it is more convenient to work this out in the form

$$\begin{aligned} \omega &= \frac{m_e c^2 \alpha^2 (1 - \frac{1}{4})}{2\hbar} = \frac{3\alpha^2}{8} \frac{1}{1.3 \times 10^{-21}} \text{ rad/s} \\ &\cong 1.5 \times 10^{16} \end{aligned}$$

Equivalently,

$$\begin{aligned} \lambda &= 2\pi \frac{c}{\omega} = \frac{16\pi}{3\alpha^2} \frac{\hbar}{m_e c} \cong 1.2 \times 10^{-7} \text{ m} \\ &= 120 \text{ nm} = 1200 \text{ \AA} \end{aligned}$$

which lies in the ultraviolet. This is in excellent agreement with experiment, and it explained some observations not previously understood. For example, certain spectral lines (known as the *Pickering series*) that mimicked the α -Lyman series ($n_1 = 1$) but for which the wavelengths were reduced in comparison with those in hydrogen by a factor of 4.0016 were explained as those associated with transitions in singly ionized helium (so that $Z = 2$, but only one electron is present), provided that account is taken of the finiteness of the nuclear mass. This changes the electron mass m_e in the formula to the *reduced mass*

electron mass $m_e(1 + m_e/M)$, where M is the nuclear mass. Nevertheless, it was clear that the theory was provisional. It said nothing about when the electrons would make their jump, nor what direction the photon of light would travel in, when emitted. Also, the quantization rules were limited to periodic systems. For elliptic orbits, the more general rule provided by A. Sommerfeld and W. Wilson stated that

$$\oint p_k dx_k = n_k h$$

where the integral ranges over a closed orbit and k ranges over the degrees of freedom; however, this was of no help in treating problems other than those associated with energy levels of hydrogenlike (one electron) atoms.

EXAMPLE 1-7

A series of lines in hydrogen correspond to transitions to a final state characterized by some quantum number n . If the wavelength of the radiation giving rise to the first line is 657 nm, what are the wavelengths corresponding to the next two lines? Assume that $\Delta n = 1$.

SOLUTION We have in general,

$$\frac{hc}{\lambda} = \frac{1}{2} m_e c^2 \alpha^2 \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right)$$

so that for our case

$$\frac{1}{\lambda} = \frac{(0.5)(0.51 \times 10^6 \text{ eV}) \left(\frac{1}{137} \right)^2}{1.24 \times 10^{-6} \text{ eV} \cdot \text{m}} \left(\frac{1}{n^2} - \frac{1}{(n+1)^2} \right)$$

that is,

$$0.152 \times 10^7 = 1.096 \times 10^7 \left(\frac{1}{n^2} - \frac{1}{(n+1)^2} \right)$$

or

$$\frac{1}{n^2} - \frac{1}{(n+1)^2} = 0.139$$

For $n = 1$ we get 0.75; for $n = 2$ we get 0.139, so $n = 2$. For the next wavelengths we get $1/\lambda = 1.096 \times 10^7(0.25 - 0.063)$, so that $\lambda = 487 \text{ nm}$; $1/\lambda = 1.096 \times 10^7(0.25 - 0.04)$, so that $\lambda = 434 \text{ nm}$.

Direct experimental evidence for the quantization of angular momentum came from an experiment performed by O. Stern and W. Gerlach in 1922. Their experimental set-up is shown in Fig. 1-13. A beam of electrically neutral silver atoms, known to possess a magnetic moment μ , was directed through an inhomogeneous magnetic field. The interaction of a magnetic moment μ with a magnetic field $\mathbf{B}$ gives rise to a change in energy.

$$V = -\mu \cdot \mathbf{B} \quad (1-41)$$

and thus a force on the atom

$$\mathbf{F} = -\nabla V = \nabla(\mu_x B_x + \mu_y B_y + \mu_z B_z) \quad (1-42)$$

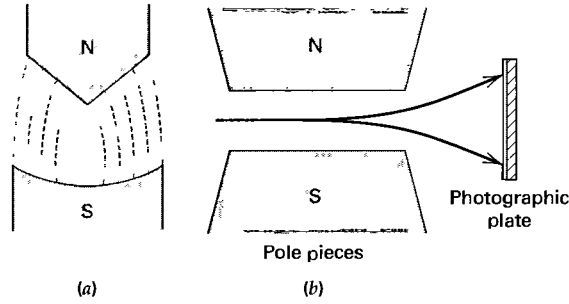


Figure 1-13 Stern-Gerlach apparatus. (a) Head-on view: The tapering of the N and S pole pieces is done so that there is an inhomogeneous magnetic field. (b) The side view shows how the magnetic moments are deflected by the field. Instead of a smear on the screen, which is expected if the magnetic moments are randomly oriented in arbitrary directions, there are only a finite number of lines, as expected from the quantization of angular momentum.

With the magnetic field primarily pointing in the z direction (as shown in Fig. 1-13) it turns out that the magnetic moment precesses about the z -direction, sufficiently rapidly that in the time it takes the atom to get through the apparatus, the average values of μ_x and μ_y vanish. Thus the force is in the z -direction, of magnitude

$$F_z = \mu_z \frac{\partial B_z}{\partial z} \quad (1-43)$$

Since μ is actually proportional to the angular momentum of the atom, one expects the z -component of the angular momentum to lie between $\pm L$, where L denotes the angular momentum. Thus, the silver atoms are subject to forces that deflect them out of the straight line along which the beam traveled initially. The experimental result that the silver atoms were deposited on the screen not in a continuous band but in two discrete lines above and below the center of the expected classical “smear” suggests strongly that angular momentum cannot take on all possible values in a certain range, but is quantized. A more detailed discussion of this can be found elsewhere.¹⁰

The Correspondence Principle

Niels Bohr made great use of the notion that the quantum theory that he had developed should merge into classical theory in the limit in which classical theory was known to apply. This idea was formulated as the *correspondence principle*. Technically it stated that the classical limit should be reached when the “quantum numbers” are large—for example, for large n in the Bohr atom. Once a consistent theory of quantum phenomena was constructed, it automatically contained classical physics as a limit, but the principle was very helpful in guiding theoretical guesses, and led Heisenberg to the point from which he could make his giant leap to quantum mechanics. To illustrate how the correspondence principle is satisfied by the Bohr atomic model, consider the frequency of the radiation emitted when an electron makes a “jump” from the orbit with quantum number $n + 1$ to the orbit with quantum number n , when n is very large. This is a good domain to ask for the classical limit, since the angular momentum $n\hbar$ is indeed much larger than $\hbar$. Classically an electron moving in a circular orbit with velocity v would be expected to radiate with the frequency of its motion; that is,

$$\nu_{cl} = \frac{v}{2\pi r} = \frac{Z\alpha c}{n} \frac{Z\alpha m_e c}{2\pi n^2 \hbar} = \frac{(Z\alpha)^2 m_e c^2}{2\pi \hbar} \frac{1}{n^3} \quad (1-44)$$

¹⁰See J. Bernstein, P. Fishbane, and S. Gasiorowicz, *loc. cit.*

On the other hand, the frequency of the radiation associated with the transition is, according to (1-23)

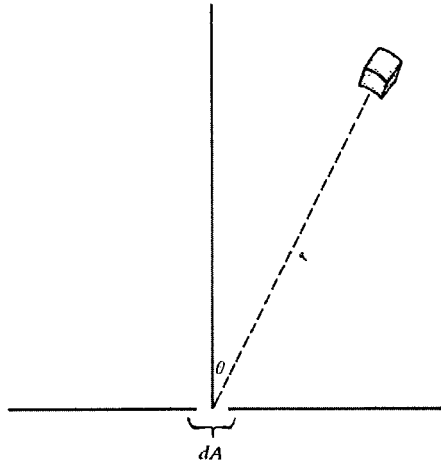
$$\nu = \frac{\omega}{2\pi} = \frac{1}{2\pi\hbar} \frac{m_e c^2 (Z\alpha)^2}{2} \left(\frac{1}{n^2} - \frac{1}{(n+1)^2} \right) \quad (1-45)$$

which approaches ν_{cl} for $n \gg 1$. Note that this is a significant result, since it is only the frequency associated with an $n + 1 \rightarrow n$ transition that corresponds to the fundamental classical frequency. The radiation associated with the jump $n + 2 \rightarrow n$ has no classical counterpart even in the large n limit. We shall see in Chapter 17 that there are no $n + 2 \rightarrow n$ transitions for “circular orbits” in quantum mechanics.

In 1925 the modern theory of quantum mechanics started with the work of Werner Heisenberg, Max Born, Pascual Jordan, Erwin Schrödinger, and Paul Dirac. This theory provides a way of reconciling all of the conflicting concepts at the cost of making us abandon a certain amount of classical thinking. It is one of the joys of being a student of physics to be able to appreciate this beautiful theory and the monumental advances in our understanding the properties of matter that the theory enabled us to make.

PROBLEMS

1. Prove the relation (1-1) between the energy density in a cavity and the emissive power. [*Hint:* To do so, look at the figure. The shaded volume element is of magnitude $r^2 dr \sin \theta d\theta d\phi = dV$, where r is the distance to the origin (at the aperture of area dA), θ is the angle with the vertical, and ϕ is the azimuthal angle about the perpendicular axis through the opening. The energy contained in the volume element is dV multiplied by the energy density. The radiation is isotropic, so that what emerges is given by the solid angle $dA \cos \theta / 4\pi r^2$ multiplied by the energy. This is to be integrated over the angles ϕ and θ and, if the flow of radiation in time Δt is wanted, over dr from 0 to $c\Delta t$ —the distance from which the radiation will escape in the given time interval.]



2. Given (1-6), calculate the energy density in a wavelength interval $\Delta\lambda$. Use your expression to calculate the value of $\lambda = \lambda_{\max}$, for which this density is maximal. Show that $\lambda_{\max}$ is of the form b/T , calculate b , and use 6000K as an estimate of the sun's surface temperature to calculate $\lambda_{\max}$ for solar radiation. [*Hint:* In calculating b you will need the solution x of the equation $(5 - x) = 5e^{-x}$.]
3. Ultraviolet light of wavelength 350 nm falls on a potassium surface. The maximum energy of the photoelectrons is 1.6 eV. What is the work function of potassium?

4. The maximum energy of photoelectrons from aluminum is 2.3 eV for radiation of 200 nm and 0.90 eV for radiation of 258 nm. Use these data to calculate Planck's constant and the work function of aluminum.
5. A 100-MeV photon collides with a proton that is at rest. What is the maximum possible energy loss for the photon?
6. A 100-keV photon collides with an electron at rest. It is scattered through 90° . What is its energy after the collision? What is the kinetic energy in eV of the electron after the collision, and what is the direction of its recoil?
7. An electron of energy 100 MeV collides with a photon of wavelength 3×10^6 nm (corresponding to the universal background of blackbody radiation). What is the maximum energy loss suffered by the electron?
8. A beam of X rays is scattered by electrons at rest. What is the energy of the X rays if the wavelength of the X rays scattered at 60° to the beam axis is $0.035 \text{ \AA}$?
9. A nitrogen nucleus (mass $\cong 14 \times$ proton mass) emits a photon of energy 6.2 MeV. If the nucleus is initially at rest, what is the recoil energy of the nucleus in eV?
10. Consider a crystal with planar spacing 0.32 nm. What order of magnitude of energies would one need for (a) electrons, (b) helium nuclei (mass $\cong 4 \times$ proton mass) to observe up to three interference maxima?
11. The smallest separation resolvable by a microscope is of the order of magnitude of the wavelength used. What energy electrons would one need in an electron microscope to resolve separations of (a) 15 nm, (b) 0.5 nm?
12. If one assumes that in a stationary state of the hydrogen atom the electron fits into a circular orbit with an integral number of wavelengths, one can reproduce the results of the Bohr theory. Work this out.
13. The distance between adjacent planes in a crystal are to be measured. If X rays of wavelength $0.5 \text{ \AA}$ are detected at an angle of 5° , what is the spacing? At what angle will the second maximum occur?
14. Use the Bohr quantization rules to calculate the energy levels for a harmonic oscillator, for which the energy is $p^2/2m + m\omega^2 r^2/2$; that is, the force is $m\omega^2 r$. Restrict yourself to circular orbits. What is the analog of the Rydberg formula? Show that the correspondence principle is satisfied for all values of the quantum number n used in quantizing the angular momentum.
15. Use the Bohr quantization rules to calculate the energy states for a potential given by

$$V(r) = V_0 \left(\frac{r}{a} \right)^k$$

with k very large. Sketch the form of the potential and show that the energy values approach $E_n \approx Cn^2$.

16. The classical energy of a plane rotator is given by

$$E = L^2/2I$$

where L is the angular momentum and I is the moment of inertia. Apply the Bohr quantization rules to obtain the energy levels of the rotator. If the Bohr frequency condition is assumed for the radiation in transitions from states labeled by n_1 to states labeled by n_2 , show that (a) the correspondence principle holds, and (b) that it implies that only transitions $\Delta n = \pm 1$ should occur.

17. Molecules sometimes behave like rotators. If rotational spectra are characterized by radiation of wavelength of order 10^6 nm, and this is used to estimate interatomic distances in a molecule like H_2 , what kind of separations (in nm) are obtained?

Supplement 1-A

Einstein's Approach to Planck's Law

In 1917 Albert Einstein wrote a remarkable paper in which he used classical statistical mechanics and elements of the old Bohr theory to derive the Planck distribution and to relate spontaneous emission, as it appears in the Bohr theory to induced emission and absorption.

We start with radiation in a cavity at temperature T . Assume that the walls of the cavity consist of atoms that can be in a variety of stationary states. These are labeled by energies E_n , and furthermore we allow for *degeneracy*—that is, the possibility that each energy level may accommodate g_n states. According to classical statistical mechanics (Boltzmann distribution), the ratio of the number of atoms in the state m to those in the state n is given by

$$\frac{N_m}{N_n} = \frac{g_m e^{-E_m/kT}}{g_n e^{-E_n/kT}} = \frac{g_m}{g_n} e^{-(E_m - E_n)/kT} \quad (1A-1)$$

Let us now consider a pair of levels of energies E_1 and E_0 with $E_1 > E_0$ (Fig. 1A-1). The number of transitions from the upper to the lower state per second—that is, the *rate* of transitions—is denoted by R_{10} . The rate of transitions from the lower to the upper level is R_{01} . The number of transitions from the lower level to the upper level must be proportional to the number of atoms in the lower state N_0 , as well as to the intensity of the radiation in the cavity. For frequency ν and temperature T this is denoted by $u(\nu, T)$. We write

$$R_{01} = N_0 u(\nu, T) B_{01} \quad (1A-2)$$

The frequency ν is the one that is absorbed in the *induced absorption*, and the constant B_{01} is called the coefficient of induced absorption.

For the transition from the upper level to the lower level, Einstein uses the Bohr postulate, according to which spontaneous emission occurs with a rate independent of the radiation present. In addition, the presence of radiation also induces transitions to the lower level. The coefficient of *induced emission* is denoted by B_{10} . If the number of atoms in the upper state is N_1 , then

$$R_{10} = N_1 (A_{10} + u(\nu, T) B_{10}) \quad (1A-3)$$

with the two terms describing spontaneous and induced emission. In equilibrium there must be as many “up” transitions as “down” transitions, so that

$$R_{10} = R_{01} \quad (1A-4)$$

It follows that

$$\frac{N_0}{N_1} = \frac{A_{10} + B_{10} u(\nu, T)}{B_{01} u(\nu, T)} = \frac{g_0}{g_1} e^{-(E_0 - E_1)/kT} \quad (1A-5)$$

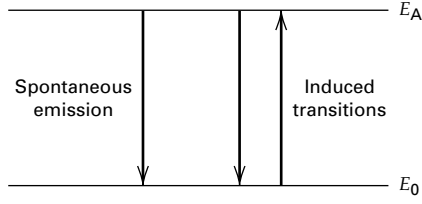


Figure 1A-1 Possible transitions in a two-level system. Downward transitions can be both spontaneous and induced by an external electromagnetic field. Upward transitions can only be stimulated.

We rewrite this in the form

$$g_1 A_{10} = u(\nu, T)(g_0 B_{01} e^{(E_1 - E_0)/kT} - g_1 B_{10}) \quad (1A-6)$$

A number of consequences can be drawn from this formula. First, in the limit of large T , for fixed $(E_1 - E_0)$,

$$e^{(E_1 - E_0)/kT} \rightarrow 1 + \frac{E_1 - E_0}{kT} + 0\left(\frac{1}{(kT)^2}\right) \quad (1A-7)$$

Furthermore, in that limit, the classical Rayleigh-Jeans formula for the energy density applies, so that

$$u(\nu, T) \rightarrow \frac{8\pi\nu^2}{c^3} kT \quad (1A-8)$$

The left side of (1A-6) is independent of T , so it follows that the term proportional to T on the right-hand side must vanish. This implies

$$g_0 B_{01} = g_1 B_{10} \quad (1A-9)$$

The meaning of this result is that the rate per atom of induced absorption is equal to the rate per atom of induced emission. The equality of the terms independent of T leads to

$$\frac{A_{10}}{B_{10}} = \frac{8\pi\nu^2}{c^3} (E_1 - E_0) \quad (1A-10)$$

If we now go back to equation (1A-6), we find that

$$u(\nu, T) = \frac{A_{10}/B_{10}}{e^{(E_1 - E_0)/kT} - 1} = \frac{8\pi\nu^2}{c^3} \frac{(E_1 - E_0)}{e^{(E_1 - E_0)/kT} - 1} \quad (1A-11)$$

The left-hand side must, on general thermodynamic grounds, obey Wien's law; that is, it must be of the form $\nu^3 f(\nu/T)$. This then has the following important consequences: $(E_1 - E_0)$ has to be proportional to ν , so that

$$E_1 - E_0 = h\nu \quad (1A-12)$$

and finally,

$$u(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1} = \left(\frac{8\pi\nu^2}{c^3}\right) h\nu \frac{1}{e^{h\nu/kT} - 1} \quad (1A-13)$$

This expression for the energy density has three components: the first one is the number of modes for the radiation field in the vicinity of ν . More precisely, we should be talking about $u(\nu, T)d\nu$; the energy density in the frequency range $(\nu, \nu + d\nu)$, and the corresponding number of modes is $(8\pi\nu^2 d\nu/c^3)$. The second factor is the energy per mode; we associate it with the energy of the quantum of radiation emitted by the decay of the upper to the lower state, because of energy conservation. Finally, the third factor is the

average number of quanta with frequency ν that make up the energy density, so that we can write

$$\langle n(\nu) \rangle = \frac{1}{e^{h\nu/kT} - 1} \quad (1A-14)$$

Let us now get back to the *emission* rate per atom. It is

$$\begin{aligned} R_{10}/N_1 &= u(\nu, T)B_{10} + A_{10} = A_{10} \left(1 + \frac{1}{e^{h\nu/kT} - 1} \right) \\ &= A_{10}(1 + \langle n(\nu) \rangle) \end{aligned} \quad (1A-15)$$

The *absorption* rate per atom is similarly found to be proportional to $\langle n(\nu) \rangle$. In *photon* language we see that both emission and absorption are enhanced by the number of photons of the right frequency that are present. Although this result was only derived for blackbody radiation, this result does not depend on the particular form of the frequency distribution.

Supplement 1-B

Estimate of the Lifetime of a Rutherford Atom

We begin with the form of $F = ma$ for a circular orbit of an electron in the Coulomb field of a proton. The equation reads

$$\frac{mv^2}{r} = m\omega^2 r = \frac{e^2}{4\pi\epsilon_0 r^2} \quad (1B-1)$$

which leads to Kepler's third law

$$\omega^2 r^3 = \frac{e^2}{4\pi\epsilon_0 m} \quad (1B-2)$$

The energy can be expressed in terms of r with the help of the above:

$$\begin{aligned} E &= \frac{1}{2} mv^2 - \frac{e^2}{4\pi\epsilon_0 r} = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0 r^2} r - \frac{e^2}{4\pi\epsilon_0 r} \\ &= -\frac{1}{2} \frac{e^2}{4\pi\epsilon_0 r} \end{aligned} \quad (1B-3)$$

The Larmor formula for the instantaneous power radiated—that is, the energy loss per unit time—is

$$\begin{aligned} P = \frac{dE}{dt} &= -\frac{2}{3} \frac{e^2}{4\pi\epsilon_0} \frac{a^2}{c^3} = -\frac{2}{3} \frac{e^2}{4\pi\epsilon_0} \frac{1}{c^3} \left(\frac{v^2}{r} \right)^2 = -\frac{2}{3} \frac{e^2}{4\pi\epsilon_0} \frac{1}{c^3} (\omega^2 r)^2 \\ &= -\frac{2}{3} \frac{e^2}{4\pi\epsilon_0} \frac{1}{c^3} \frac{1}{r^4} (\omega^2 r^3)^2 = -\frac{2}{3} \frac{e^2}{4\pi\epsilon_0} \frac{1}{c^3} \left(\frac{e^2}{4\pi\epsilon_0 m} \right)^2 \frac{1}{r^4} = -\frac{2}{3} \left(\frac{e^2}{4\pi\epsilon_0} \right)^3 \frac{1}{m^2 c^3 r^4} \end{aligned} \quad (1B-4)$$

On the other hand, it follows from the expression for the energy that

$$\frac{dE}{dt} = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \frac{1}{r^2} \frac{dr}{dt} \quad (1B-5)$$

If we combine these equations we get

$$3r^2 \frac{dr}{dt} = -4 \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{1}{m^2 c^3} \quad (1B-6)$$

This may be integrated, and it leads to

$$4 \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{1}{m^2 c^3} T = [r^3(t=0) - r^3(t=T)] \quad (1B-7)$$

It simplifies matters if we write

$$\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 = \left(\frac{e^2}{4\pi\epsilon_0\hbar c}\right)^2 \hbar^2 c^2 = \alpha^2 \hbar^2 c^2 \quad (1B-8)$$

We take $r(T) = 0$, and for $r(0)$ we take a number of the order of $10^{-10} m$. We know from our discussion of the Bohr atom that this is approximately equal to $2a_0$ where a_0 is the Bohr radius. When we put all this together we get

$$T = \left(2 \frac{\hbar}{mc\alpha}\right)^3 \frac{m^2 c}{4\alpha^2 \hbar^2} = 2 \frac{\hbar}{mc^2} \frac{1}{\alpha^5} \quad (1B-9)$$

If we put in numbers we get $T = 1.3 \times 10^{-10}$ sec.

Chapter 2

Wave-Particle Duality, Probability, and the Schrödinger Equation

The developments outlined in Chapter 1 are often described as the *Old Quantum Theory*. The rules devised were all *ad hoc*, and the connection between various separate discoveries, such as the particle nature of radiation, the wave nature of electrons and the Bohr atom (as well as other rules not discussed in our brief survey) did not rest on any firm foundation. Quantum mechanics was discovered twice: first, by Werner Heisenberg in 1925 as matrix mechanics, and then again by Erwin Schrödinger in 1926 as wave mechanics. The two forms were soon found to be identical in content, but wave mechanics became a more useful tool because the mathematics of waves were familiar to many physicists. In this chapter we begin our study of quantum mechanics, and we follow, in spirit, the path laid out by Schrödinger.

2-1 RADIATION AS PARTICLES, ELECTRONS AS WAVES

The fact that radiation and electrons exhibit both particle and wave properties raises deep conceptual difficulties, as can be seen from the following considerations: There is no doubt that light consists of individual particles, called *photons*, which carry energy and momentum, as was first unequivocally demonstrated by the Compton effect. The human eye cannot detect individual photons, but it is fairly close to being a photon counter, since under optimal conditions it takes only 5–10 photons to activate the darkness-adapted eye. There are devices, known as *photomultipliers*, that can easily detect individual photons.

An interesting *thought experiment* is discussed in Dirac's wonderful book on quantum mechanics. When light of a certain polarization is used to produce electrons (as in the photoelectric effect), the latter are emitted with an angular distribution that depends on the direction of the polarization of the incident photon beam. Since in the photoelectric effect single photons eject single electrons, this implies that individual photons, in addition to carrying energy and momentum, also have polarization properties. Suppose we now send a beam of polarized light with an initial intensity I_0 through a crystal, which has the property that only the component of light polarized along a particular axis can pass through it. Thus if the polarization of the initial beam is in the direction of the axis, then the emerging beam will have intensity I_0 . If the polarization vector makes an angle θ with the axis, then the intensity of the emerging beam is $I_0 \cos^2 \theta$. Let us look at this result in terms of individual photons. If the beam is *totally* polarized along the direction of the axis, then all

of the photons that make up the beam must have been polarized in that direction. For a beam polarized in a different direction, the intensity is reduced by a factor $\cos^2 \theta$. This implies that only this fraction of the photons passes through the crystal. However, *photons cannot be split into pieces*, so that a given photon will either pass through the crystal, or it will not. We have no way of predicting whether an individual photon will pass through. All we can say is that for N incident photons, $N \cos^2 \theta$ will get through, so that the odds, or *probability*, that a particular photon will get through is $\cos^2 \theta$.

We also know from classical optics that a beam of light consisting of many photons will exhibit wavelike properties—that is, diffraction and interference. An experiment carried out by G. I. Taylor in 1909 was the first to show that a beam of light gave rise to a diffraction pattern around a needle even when the intensity of the light was so low that only one photon at a time passed by the needle. Since then, many more experiments showed that the interference and diffraction properties cannot be due to the collective effect of the many photons in a beam. This raises new problems. Consider a thought experiment, which is a variant of the Taylor experiment, in which a very low intensity beam of light is directed at a screen with two slits in it. The photons are then detected at a second screen (Fig. 2-1). The intensity is such that at a given time no more than one photon passes through the two-slit screen. After very many photons have passed by, we see the classically expected diffraction pattern. Classically this is well understood: If the electric fields at a particular point $\mathbf{r}$ on the detecting screen due to electromagnetic waves crossing slits 1 and 2 are $\mathbf{E}_1(\mathbf{r}, t)$ and $\mathbf{E}_2(\mathbf{r}, t)$ respectively, then the total field at the point $\mathbf{r}$ at the time t is the *sum of the fields*. This is a consequence of the superposition rules for electric fields, which in turn is a consequence of the fact that Maxwell's equations for the electromagnetic fields are *linear*. The intensity at the screen is proportional to the square of the total electric field, and thus to $(\mathbf{E}_1(\mathbf{r}, t) + \mathbf{E}_2(\mathbf{r}, t))^2$. The interference pattern is due to the presence of the $\mathbf{E}_1(\mathbf{r}, t) \cdot \mathbf{E}_2(\mathbf{r}, t)$ cross term in the square of the sum of the fields. If only slit 1 were open, the intensity would be proportional to $\mathbf{E}_1(\mathbf{r}, t)^2$, and if only slit 2 were open, the intensity would be proportional to $\mathbf{E}_2(\mathbf{r}, t)^2$. If we now translate *intensity* into *probability*, as suggested by our discussion about polarization, we find that if only slit 1 is open, the probability of finding a photon at $\mathbf{r}$ is $P_1(\mathbf{r}, t)$, and if only slit 2 is open, the probability of finding a photon at $\mathbf{r}$ is $P_2(\mathbf{r}, t)$. However, if *both slits are open*, the probability is *not* the sum of the probabilities associated with each slit.

The only way to resolve these difficulties is to assume that *each photon interferes with itself*. This can be handled by assuming that each photon is described by its own electric field, $\mathbf{e}(\mathbf{r}, t)$, and that in the presence of two slits, the photon field at the detector is the sum of two terms. These are associated with the *presence of two slits*, so that

$$\mathbf{e}(\mathbf{r}, t) = \mathbf{e}_1(\mathbf{r}, t) + \mathbf{e}_2(\mathbf{r}, t)$$

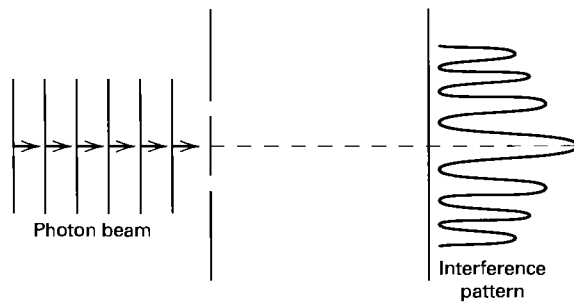


Figure 2-1 Interference pattern resulting from the passage of a beam of photons through a screen with two open slits.

just as for a classical light wave. Note that we are still talking about a *single photon*. The only real requirements are (1) that the field $\mathbf{e}(\mathbf{r}, t)$ obeys a *linear* equation and that (2) in the classical limit a large collection of photons acts in accordance with Maxwell's equations. The actual formulation of a quantum theory of photons is somewhat complicated, and we leave the discussion of this to Supplement 18-A.

At this point we turn our attention to electrons. From the time of their discovery, electrons were described as *particles*. They appear to travel along trajectories determined by the electric and magnetic forces acting on them, they have mass, and they carry energy and momentum. Nevertheless, they have wavelike properties, as first determined in the diffraction of electron beams by crystals. As demonstrated by the beautiful *two-slit* experiment of A. Tonomura presented in Chapter 1 (see Fig. 1-9) the pattern of the hits of electrons on the absorbing screen slowly builds up to a wavelike interference pattern. This experiment is a realization of the *thought experiments* we discussed above in connection with photons. Individual electrons appear to hit the screen at random. As the number of electrons increases, the expected interference pattern emerges. Again, the place where a single electron hits cannot be affected by the fact that other electrons came before it, or will come after it, so that the emergence of the pattern must lie in the property of each electron. By analogy of our conjectures in connection with the photon, we may expect that the properties of a single electron are described by an analog of the one-photon electric field $\mathbf{e}(\mathbf{r}, t)$. We thus expect that an electron will be described by a *wave function* traditionally denoted by $\psi(\mathbf{r}, t)$. In order to get an electron to interfere with itself we must insist that $\psi(\mathbf{r}, t)$ obeys a linear equation. In that case, a sum of two wave functions is also a wave function, so that the superposition rules apply. Furthermore, we expect that the predictability of where an electron hits (as if it were a classical particle) will be replaced by a statement, involving the wave function $\psi(\mathbf{r}, t)$, about the probability that an electron arrives at $\mathbf{r}$. The rest of this chapter is devoted to arguments that lead us to the correct form of the linear equation obeyed by $\psi(\mathbf{r}, t)$, the **Schrödinger equation**, the probability for finding an electron at $\mathbf{r}$ at a time t in terms of $\psi(\mathbf{r}, t)$ and other general properties of the wave function. We shall approach this by constructing waves that might simulate the properties of particles. The study of *wave packets* will be helpful in this matter, even though the idea that there are real waves that act like particles is not correct.

2-2 PLANE WAVES AND WAVE PACKETS

A harmonic wave propagating in the positive x -direction with wave number k has the form

$$\psi_k(x, t) = A_1 \cos(kx - \omega t) + A_2 \sin(kx - \omega t)$$

or equivalently

$$\psi_k(x, t) = Ae^{i(kx - \omega t)} + Be^{-i(kx - \omega t)} \quad (2-1)$$

The *wave number* k is related to the wavelength by

$$k = 2\pi/\lambda \quad (2-2)$$

and the *angular frequency* ω is related to the period T by

$$\omega = 2\pi/T$$

and thus to the frequency $\nu = 1/T$ by

$$\omega = 2\pi\nu \quad (2-3)$$

In general ω will be related to k in some way. For example, in the case of light propagating in a vacuum, $\nu = c/\lambda$ where c is the speed of light. Hence $\omega = kc$. This relationship is *not* true for light in a dispersive medium. There $\nu = c/n\lambda$, where n is the refractive index of the medium, and n is generally a function of the wavelength $n = n(\lambda)$. We shall determine below the relation between ω and k for the waves of interest to us.

Since $\psi_k(x, t)$ does not depend on y or z , it takes on the same value everywhere on the y - z plane, and is thus called a *plane wave*. $\psi_k(x, t)$ is a plane wave for all possible values of k , and it is therefore possible to take a *superposition* of plane waves with different amplitudes $A(k)$ and $B(k)$. There is nothing in what we have done so far to require $A(k)$ and $B(k)$ to be real, and, in fact, we shall see that in general we must allow $\psi(x, t)$ to be complex. Let us consider the wave $A(k)e^{i(kx - \omega t)}$ and add such waves for a variety of values of k .¹ The superposition is called a *wave packet*, and it takes the form

$$\psi(x, t) = \int_{-\infty}^{\infty} dk A(k) e^{i(kx - \omega t)} \quad (2-4)$$

We begin by considering the wave packet at time $t = 0$,

$$\psi(x, 0) = \int_{-\infty}^{\infty} dk A(k) e^{ikx} \quad (2-5)$$

and illustrate it by considering a special form, called the *gaussian* form

$$A(k) = e^{-\alpha(k - k_0)^2/2} \quad (2-6)$$

This function is centered about k_0 , and it falls off rapidly away from that center. We shall see that it is the *width* of the square of this function that is of interest. The square falls to 1/3 of its peak value when $\alpha(k - k_0)^2 \cong 1$. This means that the *width* can be taken to be $\Delta k = 2/\sqrt{\alpha}$. The integral in (2-5) can now be done in steps. We first make a change of variables to $q' = k - k_0$, and we end up with

$$\begin{aligned} \psi(x, 0) &= e^{ik_0 x} e^{-x^2/2\alpha} \int_{-\infty}^{\infty} dq' e^{-\alpha q'^2/2} \\ &= \sqrt{\frac{2\pi}{\alpha}} e^{ik_0 x} e^{-x^2/2\alpha} \end{aligned} \quad (2-7)$$

Aside from the scale factor, which could have been absorbed by slightly modifying $A(k)$, we end up with a plane-wave factor characterized by the wave number k_0 and a modulating function that acts to *localize* the packet about $x = 0$. The width of that packet, which is also *gaussian*, is defined in the same way as before: We square the function and see where it drops off to about 1/3. This provides the width $\Delta x = 2\sqrt{\alpha}$. We see that there is a reciprocal relation between the width of the function $A(k)$ that determines the shape of the wave packet, and the width of the wave packet. In fact, the product of the two widths has the property that $\Delta k \Delta x = 4$ here. The specific value of the number on the right side is not important. A slightly different definition of the widths would have given us a different an-

¹We do not consider the wave form $B(k) \exp[-i(kx - \omega t)]$ because a term proportional to $\exp[i\omega t]$ will be seen to be associated with a negative kinetic energy.

swer. What *is* quite general is that the product is *independent of α* . It is actually a general result of Fourier integrals that

$$\Delta k \Delta x > \frac{1}{2} \quad (2-8)$$

so that this reciprocal relation is true for wave packets in general. We illustrate this in the example that follows.

EXAMPLE 2-1

Consider a wave packet for which

$$\begin{aligned} A(k) &= N & -K \leq k \leq K \\ &= 0 & \text{elsewhere} \end{aligned}$$

Calculate $\psi(x, 0)$, and use some reasonable definition of the width to show that (2-8) is satisfied.

SOLUTION We have

$$\psi(x, 0) = \int_{-K}^K dk N e^{ikx} = \frac{N}{ix} (e^{iKx} - e^{-iKx}) = 2N \frac{\sin Kx}{x}$$

The definition of $A(k)$ easily shows that $\Delta k = 2K$. A reasonable definition of Δx might be the distance between the two points at which $\psi(x)$ first vanishes as it gets away from $x = 0$. This happens when $Kx = \pm\pi$, so that $\Delta x = 2\pi/K$. It follows that

$$\Delta k \Delta x = 4\pi$$

which certainly satisfies (2-8).

The mathematical description of how wave packets move is a little messy, so we set the material off in a subsection. The important result is that for a wave packet for which the spread about a particular value of k —say, k_0 —is small, so that it looks a lot like a plane wave modulated by a very wide function in x , then the wave packet moves with the *group velocity*

$$v_g = \left(\frac{\partial \omega(k)}{\partial k} \right)_{k=k_0} \quad (2-9)$$

and its width in x spreads as a function of time. In the following subsection we see this in detail for our Gaussian wave packet.

*How Wave Packets Move

As indicated in (2-4), the motion of the wave package can be obtained by integration, *provided we know how ω depends on k* . Let us assume that $A(k)$ is sharply peaked about the value of $k = k_0$. We may make the approximation

$$\omega(k) = \omega(k_0) + (k - k_0) \left(\frac{\partial \omega}{\partial k} \right)_{k=k_0} + \frac{1}{2} (k - k_0)^2 \left(\frac{\partial^2 \omega}{\partial k^2} \right)_{k=k_0} \quad (2-10)$$

Taylor expansion

With this,

$$(kx - \omega t) = (k_0 x - \omega(k_0)t) + (k - k_0) \left[x - \left(\frac{\partial \omega}{\partial k} \right)_{k_0} t \right] - \frac{1}{2} (k - k_0)^2 \left(\frac{\partial^2 \omega}{\partial k^2} \right)_{k_0} t$$

When this is inserted into (2-4) we get, after changing variables to $q = k - k_0$, the following:

$$\psi(x, t) = e^{i(k_0 x - \omega(k_0)t)} \int_{-\infty}^{\infty} dq A(q + k_0) e^{iq(x - v_g t)} e^{-iq^2(\partial^2 \omega / \partial k^2)_0 t / 2} \quad (2-11)$$

If the last factor in the integrand were absent, as is the case when $\omega = kc$ for example, then the integral becomes a function of $(x - v_g t)$. If this function peaks at $x = 0$ at time $t = 0$, then it will peak at $x = v_g t$ at a later time t . The velocity with which the packet moves is the *group velocity*. The second term in the integrand, when it does not vanish, modifies the amplitude $A(q + k_0)$ and thus the shape of the wave packet; this is best seen in terms of our gaussian packet. The integral is worked out just like the one that leads to (2-7), since the extra term is also gaussian, as the exponent is quadratic in q^2 . With the notation,

$$\left(\frac{\partial^2 \omega}{\partial k^2} \right)_{k=k_0} = \beta$$

We see that in addition to changing x to $x - v_g t$, we change α to $\alpha + 2i\beta t$. We therefore get

$$\psi(x, t) = \sqrt{\frac{2\pi}{\alpha + 2i\beta t}} e^{i(k_0 x - \omega_0 t)} e^{-\frac{(x - v_g t)^2}{2\alpha + 4i\beta t}} \quad (2-12)$$

This is a rather untransparent expression, involving a complex function of x and t , but as we shall see soon, its absolute magnitude has a physical meaning. We therefore calculate

$$|\psi(x, t)|^2 = \psi^*(x, t)\psi(x, t)$$

The pre-factor is easily handled; the phase factor has magnitude 1, and in the second exponential, the exponent has to be added to its complex conjugate. When all of this is done, we get

$$|\psi(x, t)|^2 = \frac{2\pi}{\sqrt{\alpha^2 + 4\beta^2 t^2}} e^{-\frac{\alpha(x - v_g t)^2}{\alpha^2 + 4\beta^2 t^2}} \quad (2-13)$$

Comparison with the value at $t = 0$ shows that the width, initially given by $2\sqrt{\alpha}$, now becomes $2\sqrt{\alpha + 4\beta^2 t^2/\alpha} = 2\sqrt{\alpha}\sqrt{1 + 4\beta^2 t^2/\alpha^2}$. The result is that the wave packet traveling with speed v_g *spreads* as time increases. When α is large, so that the packet is broad, the spreading is small. Nevertheless it is there, and this stands in the way of interpreting the wave packet as describing the particle itself.

2-3 THE PROBABILITY INTERPRETATION OF THE WAVE FUNCTION

Electric Field

At this point we recall that in the case of photons, the intensity, proportional to $[\mathbf{e}(\mathbf{r}, t)]^2$, was interpreted as being proportional to *the probability of finding a photon in the vicinity of $\mathbf{r}$ at the time t* . Since we were led to the conclusion that $\psi(\mathbf{r}, t)$ had to be complex, we assume that it is $|\psi(\mathbf{r}, t)|^2$ that is related to the corresponding probability of finding an electron in the vicinity of $\mathbf{r}$ at time t . For simplicity we deal with motion in one dimension (though the generalization is straightforward) and assert

The probability of finding an electron, described by the wave function $\psi(x, t)$, in the region lying between x and $x + dx$ is given by

$$P(x, t)dx = |\psi(x, t)|^2 dx \quad (2-14)$$

The probability interpretation is due to Max Born who, shortly after the discovery of the Schrödinger equation, studied the scattering of a beam of electrons by a target and was led to the above form.

With this interpretation of $|\psi(x, t)|^2$ the spreading of the wave packet presents no problems. All it implies is that an electron, known to be in a certain region with some probability distribution, will, with increasing time, have an increasing probability of being found outside that region.

The appearance of probability in quantum mechanics differs from its appearance in classical physics. Here it is not a statement of ignorance about what is “really” going on, as is the case when we speak of the probability of a coin-toss leading to heads or tails, but it is a basic limitation on what we can know when the wave function is known. The mathematical implications of this interpretation of $\psi(x, t)$ will be discussed at length in the next chapter.

The probability interpretation allows us to understand electron interference. As a consequence of the linearity of the equation for $\psi(x, t)$, a wave function of the form

$$\psi(x, t) = N(\psi_1(x, t) + \psi_2(x, t)) \quad (2-15)$$

is a solution, if both ψ_1 and ψ_2 are solutions. Let ψ_1 be the wave function of the electron that describes the system with slit 2 closed. This wave function is then definitely associated with passage through slit 1. Similarly, if ψ_2 is the wave function with slit 1 closed, then the wave function with both slits open will then be the sum of the wave functions ψ_1 and ψ_2 . Consequently, the probability density of finding an electron at a point x on the photographic plate behind the slits is proportional to $|\psi_1(x, t) + \psi_2(x, t)|^2$.

We have

$$|\psi_1(x, t) + \psi_2(x, t)|^2 = |\psi_1(x, t)|^2 + |\psi_2(x, t)|^2 + 2 \operatorname{Re}(\psi_1(x, t)\psi_2^*(x, t)) \quad (2-16)$$

and the third term clearly exhibits the interference. This effect requires that there be a single electron source, so that the phase difference of the two wave functions ψ_1 and ψ_2 does not vary randomly. If the phase difference varied unpredictably, then the probability would be determined by

$$P(x, t) = |\psi_1(x, t)|^2 + |\psi_2(x, t)|^2 \quad (2-17)$$

EXAMPLE 2-2

Consider a two-slit experiment, in which the wave function at slit 1 acquires an arbitrary *random* phase that is to be averaged over, so that the total wave function at the screen is $\psi(x, t) = e^{i\phi}\psi_1(x, t) + \psi_2(x, t)$. (Such a situation might arise if there were *two incoherent* electron sources, one at each slit). Show that under these circumstances the interference averages out.

SOLUTION We need to calculate

$$\begin{aligned} |\psi(x, t)|^2 &= (e^{i\phi}\psi_1(x, t) + \psi_2(x, t))(e^{-i\phi}\psi_1^*(x, t) + \psi_2^*(x, t)) \\ &= |\psi_1(x, t)|^2 + |\psi_2(x, t)|^2 + e^{i\phi}\psi_1(x, t)\psi_2^*(x, t) + e^{-i\phi}\psi_2(x, t)\psi_1^*(x, t) \\ &= |\psi_1(x, t)|^2 + |\psi_2(x, t)|^2 + 2 \cos \phi \operatorname{Re}(\psi_1(x, t)\psi_2^*(x, t)) + 2 \sin \phi \operatorname{Im}(\psi_1(x, t)\psi_2^*(x, t)) \end{aligned}$$

The angle ϕ varies randomly from electron to electron, so that in the pattern of dots, the terms involving this angle average out to zero.

There is a potential difficulty with the probability interpretation. Consider a beam of electrons passing through a screen with a double slit. Suppose we could determine in some

way through which slit each given electron passes. If the electrons are far apart, then as far as any given electron is concerned, we might as well have closed the other slit. If such a detection means were available to us, we could divide all the electrons, arriving at the screen on which they are detected, into two classes: those that went through slit 1 and those that went through slit 2. In that case, however, we would get the distribution (2-17). We are forced to the conclusion that we get an interference pattern only if the experiment does *not* allow us to determine which slit the electrons go through. If we somehow arrange to find out what the slit of passage was, then the interference pattern disappears and the probability is just the sum of the individual probabilities. The rule is simple:

If the paths are not determined, add the wave functions and square; if the paths are determined, square the wave functions and add.

If we look at the discussion in Example 2-2, we see that somehow the acquisition of knowledge in which way the electron went must introduce a random phase into the component of the wave function that is being “looked at.” For further discussion see Chapter 20.

2-4 THE SCHRÖDINGER EQUATION

We have constructed a wave function that may be used to describe satisfactorily the probability of finding a freely traveling electron at x , at a time t . We make the connection with physics by first recalling that according to de Broglie $k = p/\hbar$, and as suggested by the Planck relation, $\omega = E/\hbar$. Thus the wave packet may be rewritten in the form

$$\psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \phi(p) e^{i(px - Et)/\hbar} \quad (2-18)$$

We now take $E(p) = p^2/2m$ for a free particle. We see that the group velocity is

$$v_g = \frac{\partial \omega}{\partial k} = \frac{\partial E}{\partial p} = \frac{p}{m} \quad \text{Relation between E and G.velocity}$$

which confirms our association of $\hbar\omega$ with the energy. Here $\phi(p)/\sqrt{2\pi\hbar}$ plays the role played by $A(k)$ in eq. (2-5).

Now suppose that the particle under consideration is not free, so that instead of $E = p^2/2m$, we have

$$E = p^2/2m + V(x) \quad (2-19)$$

If we were to mindlessly insert this into the exponent in eq. (2-18), the wave function would be changed rather trivially: the new wave function $\psi(x, t)$ is just the product of the free particle wave function, and a factor $e^{iV(x)t/\hbar}$. Now this factor is a pure phase factor, whose absolute square is 1. This would mean that the addition of a potential to the energy in the exponent does not change the uniform motion of the wave packet. This is patently wrong since in a potential, the velocity of a particle changes from place to place. What we need to do is to find the equation that $\psi(x, t)$ is a solution of, and then modify that equation to take into account the presence of a potential $V(x)$.

We proceed as follows: If we differentiate with respect to time, we find that

$$\begin{aligned} i\hbar \frac{\partial \psi(x, t)}{\partial t} &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \phi(p) E(p) e^{i(px - Et)/\hbar} \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \phi(p) \frac{p^2}{2m} e^{i(px - Et)/\hbar} \end{aligned} \quad (2-20)$$

On the other hand,

$$\frac{\hbar}{i} \frac{\partial}{\partial x} \psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \phi(p) p e^{i(px-Et)/\hbar}$$

and hence

$$\left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right)^2 \psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \phi(p) p^2 e^{i(px-Et)/\hbar} \quad (2-21)$$

We may combine these two results to get an equation for $\psi(x, t)$ that is solved by (2-18). This is

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} \quad (2-22)$$

This is the **Schrödinger equation for a free particle**. Although we started from a solution of (2-22), the equation takes precedence over the solution. It is easy to see that starting from the equation, a solution of the form $e^{i(kx+\omega t)}$ would correspond to a negative kinetic energy and that $A \cos(kx - \omega t) + B \sin(kx - \omega t)$ will only be a solution if $B = iA$. To the extent that (2-22) is a translation of $E = p^2/2m$, with E being replaced by $i\hbar \frac{\partial}{\partial t}$ and p by $-i\hbar \frac{\partial}{\partial x}$, we can generalize the energy equation in the presence of a potential $V(x)$,

$$E = \frac{p^2}{2m} + V(x)$$

to the general Schrödinger equation

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x)\psi(x, t) \quad (2-23)$$

This is the basic equation that we will be working with in much of this book.

Let us return to the free particle case. The most general solution of that equation depends on the form of $\phi(p)$, and we shall now show that this is determined by the *initial condition*—that is, by $\psi(x, 0)$. In fact, if we **set $t = 0$** in the solution (2-18), we get

$$\psi(x, 0) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \phi(p) e^{ipx/\hbar} \quad (2-24)$$

and this equation determines $\phi(p)$. The solution of (2-23) is also determined by $\psi(x, 0)$. This is in contrast to the familiar wave equation

$$\frac{\partial^2 f(x, t)}{\partial t^2} = u^2 \frac{\partial^2 f(x, t)}{\partial x^2}$$

in which both $f(x, 0)$ and $(\partial f(x, t)/\partial t)_{t=0}$ have to be specified. The difference is a consequence of the fact that the Schrödinger equation is of *first order* in t . We shall see that this is closely related to the probability interpretation of $\psi(x, t)$.

*The Relation between $\phi(p)$ and $\psi(x, 0)$

The relation between $\psi(x, 0)$ and $\phi(p)$ is obtained by noting that (2-23) is a Fourier integral and thus can be inverted. Here we make use of the properties of Fourier integrals outlined in Supplement 2-A. [www.wiley.com/college/gasiorowicz]

We multiply $\psi(x, 0)$ by $e^{-ip'x/\hbar}$ and integrate over all x . This leads to

$$\int_{-\infty}^{\infty} dx \psi(x, 0) e^{-ip'x/\hbar} = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dp \phi(p) e^{i(p-p')x/\hbar} \quad (2-25)$$

We now use $\int_{-\infty}^{\infty} dx e^{i(p-p')x/\hbar} = 2\pi\hbar\delta(p-p')$ to get $\sqrt{2\pi\hbar}\phi(p')$ on the right-hand side, so that

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi(x, 0) e^{-ipx/\hbar} \quad (2-26)$$

2-5 THE HEISENBERG UNCERTAINTY RELATIONS

Let us return to the reciprocal relation obtained in eq. (2-8),

$$\Delta k \Delta x > 1/2$$

Recalling the identification of $\hbar k$ with the momentum, this relation takes the form

$$\Delta x \Delta p \geq \hbar/2 \quad (2-27)$$

This is called the *Heisenberg uncertainty relation*, or the *Heisenberg indeterminacy relation*. It arose in the context of our discussion of the wave packets, but as we now see, it is a statement about the *wave function*. We saw that $\psi(x)$ cannot describe a particle that is both well-localized in space and has a sharp momentum. This is in great contrast to classical mechanics. What the relation states is that there is a quantitative limitation on the accuracy with which we can describe a system using our familiar, classical notions of position and momentum. Position and momentum are said to be *complementary* variables. We can illustrate this limitation by a couple of examples.²

Diffraction of a Photon Beam

Consider a beam of photons, passing through a slit of width a (Fig. 2-2). When the beam is treated as an electromagnetic wave, one can show that the beam is *diffracted* with an angular spread of magnitude

$$\theta \approx \frac{\lambda}{a}$$

where λ is the wavelength of the light. The particle-wave duality of quantum mechanics allows us to describe the beam as a sequence of photons passing through the slit. Under those circumstances, one would not expect a spreading of the beam. The uncertainty relation rescues us, so to speak, from a paradox. The photon momentum in the x -direction is given by $p_x = h/\lambda$. The y -coordinate of the photon is indeterminate to the extent that

$$\Delta y \leq a$$

²Many examples are discussed in W. Heisenberg, *The Physical Principles of the Quantum Theory*, Dover Publications, New York, 1930, and in most books of quantum mechanics listed in the bibliography. See also J. Bernstein, P. M. Fishbane, and S. Gasiorowicz, *loc.cit.*

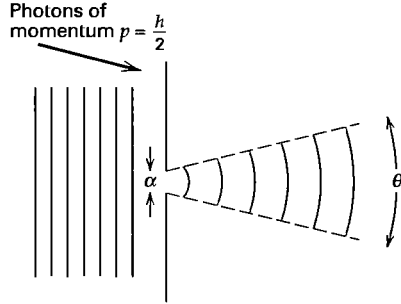


Figure 2-2 Diffraction of a beam of photons that pass through a slit in a screen.

This implies that the momentum of the photon in the y -direction is uncertain to the extent that

$$\Delta p_y \geq \frac{h}{\Delta y} > \frac{h}{a}$$

This, however corresponds to a spread of the beam of the order of

$$\theta \approx \frac{\Delta p_y}{p_x} \approx \frac{h/a}{h/\lambda} \approx \frac{\lambda}{a}$$

in agreement with the “wave description” of the incoming beam.³

Inability to Localize Bohr Orbits

Quantum mechanics does not allow us to talk about classical orbits of an electron in the Coulomb field of a proton. Suppose we want to conduct an experiment to study the location of the electron in an orbit. Let us limit ourselves to circular orbits. We will want to distinguish between an orbit of radius characterized by the quantum number n and a neighboring orbit. If we use a photon beam to do this, we need a beam of wavelength such that

$$\lambda \ll r_{n+1} - r_n = \frac{\hbar}{m_e c \alpha} [(n+1)^2 - n^2] \approx \frac{h}{m_e c \alpha} n$$

(where we have again neglected factors of the order of 2π). Such a photon, by particle-wave duality, will transfer momentum of the order of

$$p_\gamma \approx \frac{h}{\lambda} \gg \frac{m_e c \alpha}{n}$$

to the electron. This means that the energy transfer to the electron is

$$\Delta E \approx \frac{p \Delta p}{m_e} \approx \frac{p p_\gamma}{m_e} \gg \frac{m_e (c \alpha)^2}{n^2}$$

which is large enough to kick the electron out of its “orbit.” Again, the attempt to localize the electron must be accompanied by an uncontrollable momentum transfer to it.

One more comment: The heading of this section included *Relations*. We shall see in Chapter 5 that the uncertainty relation we used above follows formally from the

³In this example, we have worked with orders of magnitude, so that we have ignored factors of 2π in the difference between h and $\hbar$.

Schrödinger equation, provided we use a suitable definition of the meaning of Δx and Δp . There is another relation that is not so directly derivable, but it is nevertheless applicable. This is a relation that is familiar in classical wave theory. A wave train that is limited in time, so that it lasts a time Δt , must have a spread in frequencies given by $\Delta \nu \geq \frac{1}{\Delta t}$, which translates into

$$\Delta E \Delta t > h \quad (2-28)$$

We shall see an example of this when we discuss the *lifetime* of an electron in an excited state and the associated *width* of the spectral line in Chapter 17.

INTERIM SUMMARY

The particle nature of radiation and electrons (and other particles such as neutrons, helium nuclei, and even complex molecules) is incompatible with the observed wave properties that these particles manifest when suitable “wave-detection” experiments are performed. The suggested solution is that the particles are described by wave functions that obey a superposition principle. These must obey a linear equation (not necessarily the familiar “wave equation”), they must be complex in general, and the wave functions must be interpreted as yielding only probability statements about the behavior of individual particles. Some good guesses lead to the linear Schrödinger equation (2-23) for the wave function $\psi(x, t)$ that describes a particle in a potential $V(x)$. We also stated the content of the Born interpretation—namely, that the wave function $\psi(x, t)$ is a *probability amplitude*, and its absolute square yields a *probability*.

2-6 MORE ON THE PROBABILITY INTERPRETATION

For the interpretation given in (2-15) to hold, we must require that

$$\int_{-\infty}^{\infty} dx P(x, t) = \int_{-\infty}^{\infty} dx |\psi(x, t)|^2 = 1 \quad (2-29)$$

since the particle must be somewhere in the range $-\infty < x < \infty$. Suppose we find a solution of the Schrödinger equation for which the integral in (2-29) is not equal to unity. As a consequence of the linearity of the Schrödinger equation, if $\psi(x, t)$ is a solution, so is $A\psi(x, t)$, and with a proper choice of A one can *normalize* the wave function $\psi(x, t)$ so it does indeed satisfy (2-29). We shall see below that all we need is that

$$\int_{-\infty}^{\infty} dx |\psi(x, 0)|^2 < \infty \quad (2-30)$$

that is, *the initial state wave functions must be square integrable*. Since we may need to deal with integrals of the type

$$\int_{-\infty}^{\infty} dx \psi^*(x, t) x^n \psi(x, t)$$

and

$$\int_{-\infty}^{\infty} dx \psi^*(x, t) \left(\frac{\partial}{\partial x} \right)^n \psi(x, t)$$

we will require that the wave functions $\psi(x, 0)$ go to zero rapidly as $x \rightarrow \pm\infty$, often faster than any power of x . We shall also require that the wave functions $\psi(x, t)$ be *continuous* in x .

The Importance of Phases

The emphasis on $|\psi(x, t)|^2$ as the physically relevant quantity might lead to the impression that the *phase* of the wave function is of no importance. If we write $\psi = Re^{i\theta}$, then indeed $|\psi|^2 = R^2$ independent of θ . However, the linearity of the equation allows us to add solutions, as in our discussion of the electron interference pattern with two slits. We see that

$$|R_1 e^{i\theta_1} + R_2 e^{i\theta_2}|^2 = R_1^2 + R_2^2 + 2R_1 R_2 \cos(\theta_1 - \theta_2)$$

depends on the relative phase. An overall phase in the total wave function can be ignored, or chosen arbitrarily for convenience.

The Probability Current

The ability to normalize the wave function is only possible if the constant A mentioned above is a constant, independent of time. We now show that this is the case. We will use (2-23) and its complex conjugate, with the explicit assumption that *the potential energy $V(x)$ is real*. Under these circumstances we have

$$-i\hbar \frac{\partial \psi^*(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi^*(x, t)}{\partial x^2} + V(x) \psi^*(x, t) \quad (2-31)$$

We may use this to calculate

$$\begin{aligned} \frac{\partial}{\partial t} P(x, t) &= \frac{\partial \psi^*}{\partial t} \psi + \psi^* \frac{\partial \psi}{\partial t} \\ &= \frac{1}{i\hbar} \left(\frac{\hbar^2}{2m} \right) \left(\frac{\partial^2 \psi^*}{\partial x^2} \psi - \psi^* \frac{\partial^2 \psi}{\partial x^2} \right) \\ &= -\frac{\partial}{\partial x} \left[\frac{\hbar}{2im} \left(\psi^* \frac{\partial \psi}{\partial x} - \frac{\partial \psi^*}{\partial x} \psi \right) \right] \end{aligned}$$

If we now define the *flux* or, equivalently, the *probability current* by

$$j(x, t) = \frac{\hbar}{2im} \left(\psi^* \frac{\partial \psi}{\partial x} - \frac{\partial \psi^*}{\partial x} \psi \right) \quad (2-32)$$

we get

$$\frac{\partial}{\partial t} P(x, t) + \frac{\partial}{\partial x} j(x, t) = 0 \quad (2-33)$$

When this is integrated over all space, we find that

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} dx P(x, t) = - \int_{-\infty}^{\infty} dx \frac{\partial}{\partial x} j(x, t) = 0 \quad (2-34)$$

The last step follows from the fact that for square integrable functions $j(x, t)$ vanishes at $\pm\infty$. Incidentally, had we allowed for discontinuities in $\psi(x, t)$ we would have been led to delta functions in the flux, and hence in the probability density, which is unacceptable in a physically observed quantity. [Eq. (2-34) then implies that A is constant.]

Equation (2-33) is a conservation law analogous to the charge conservation equation in classical electrodynamics. It expresses the fact that a change in the density in a region—say, $a \leq x \leq b$ —is compensated by a net change in flux into that region

$$\frac{d}{dt} \int_a^b dx P(x, t) = - \int_a^b dx \frac{\partial}{\partial x} j(x, t) = j(a, t) - j(b, t) \quad (2-35)$$

EXAMPLE 2-3

What is the probability current for a wave function of the form

(a) $Ae^{ikx} + Be^{-ikx}$, (b) $Ae^{-\alpha x}$, (c) $R(x)e^{iS(x)/\hbar}$?

SOLUTION

(a) With $\psi(x) = Ae^{ikx} + Be^{-ikx}$, we have $d\psi/dx = ik(Ae^{ikx} - Be^{-ikx})$ and $\psi^* = A^*e^{-ikx} + B^*e^{ikx}$, so that

$$\begin{aligned} \frac{\hbar}{2im} \left(\psi^* \frac{d\psi}{dx} - \frac{d\psi^*}{dx} \psi \right) &= \frac{\hbar}{2im} [(A^*e^{-ikx} + B^*e^{ikx})ik(Ae^{ikx} - Be^{-ikx}) - \text{c.c.}] \\ &= \frac{\hbar k}{m} (|A|^2 - |B|^2) = \frac{p}{m} (|A|^2 - |B|^2) \end{aligned}$$

(b) Here $\psi(x)$ is real, which means that $j(x, 0)$ is manifestly zero. Note that this result is true for any ψ of the form A (real function of x) where A may be a complex number.

(c) Here

$$\frac{\hbar}{2im} \left(Re^{-iS/\hbar} \left(\frac{dR}{dx} e^{iS/\hbar} + iR \frac{dS}{dx} e^{iS/\hbar} \right) - \text{c.c.} \right) = \frac{1}{m} R^2 \frac{dS}{dx}$$

These comments are sufficient to proceed with out discussion of the consequences of the probability interpretation in quantum mechanics.

2-7 EXPECTATION VALUES AND THE MOMENTUM IN WAVE MECHANICS

Given the probability density $P(x, t)$, we can calculate the expectation value of x , x^2 , or, for that matter, any function of x . (See Supplement 2-B for a discussion of expectation values) [www.wiley.com/college/gasiorowicz]. We have

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} dx f(x) P(x, t) = \int_{-\infty}^{\infty} dx \psi^*(x, t) f(x) \psi(x, t) \quad (2-36)$$

Given our assumptions about the behavior of $\psi(x, t)$ as $x \rightarrow \pm\infty$, there is no problem with the convergence of the integral. Note that we inserted $f(x)$ between $\psi^*(x, t)$ and $\psi(x, t)$. Clearly the order in which we put the three functions does not matter when we integrate. We will soon find that sometimes $f(x)$ involves derivatives with respect to x , and then the order matters. Note that $\langle f(x) \rangle$ depends on time, even if it does not have an explicit time dependence, because in its definition the wave function ψ may have a nontrivial time dependence.

The above expression does not tell us how to calculate the average of the momentum, or a function of the momentum, because we don't know what to insert between ψ^* and ψ when calculating $\langle p \rangle$, for example.

The Momentum in Wave Mechanics

We approach the calculation of $\langle p \rangle$ by noting that classically

$$p = mv = m \frac{dx}{dt}$$

so we try

$$\begin{aligned} \langle p \rangle &= m \frac{d}{dt} \langle x \rangle = m \frac{d}{dt} \int_{-\infty}^{\infty} dx \psi^*(x, t) x \psi(x, t) \\ &= m \int_{-\infty}^{\infty} dx \left(\frac{\partial \psi^*(x, t)}{\partial t} x \psi(x, t) + \psi^*(x, t) x \frac{\partial \psi(x, t)}{\partial t} \right) \end{aligned} \quad (2-37)$$

Note that there is no dx/dt under the integral sign. The only quantity that varies with time is $\psi(x, t)$, and it is this variation that gives rise to a change in $\langle x \rangle$ with time. We again use the Schrödinger equation and its complex conjugate to evaluate the above. We end up with

$$\langle p \rangle = \frac{\hbar}{2i} \int_{-\infty}^{\infty} dx \left(\frac{\partial^2 \psi^*}{\partial x^2} x \psi - \psi^* x \frac{\partial^2 \psi}{\partial x^2} \right) \quad (2-38)$$

Now

$$\begin{aligned} \frac{\partial^2 \psi^*}{\partial x^2} x \psi &= \frac{\partial}{\partial x} \left[\frac{\partial \psi^*}{\partial x} x \psi \right] - \frac{\partial \psi^*}{\partial x} \psi - \frac{\partial \psi^*}{\partial x} x \frac{\partial \psi}{\partial x} \\ &= \frac{\partial}{\partial x} \left[\frac{\partial \psi^*}{\partial x} x \psi \right] - \frac{\partial}{\partial x} (\psi^* \psi) + \psi^* \frac{\partial \psi}{\partial x} - \frac{\partial}{\partial x} \left[\psi^* x \frac{\partial \psi}{\partial x} \right] + \psi^* \frac{\partial \psi}{\partial x} + \psi^* x \frac{\partial^2 \psi}{\partial x^2} \end{aligned}$$

This means that the integrand in (2-38) has the form

$$\frac{\partial}{\partial x} \left(\frac{\partial \psi^*}{\partial x} x \psi - \psi^* x \frac{\partial \psi}{\partial x} - \psi^* \psi \right) + 2\psi^* \frac{\partial \psi}{\partial x}$$

Because the wave functions vanish at infinity, the first term does not contribute, and the integral gives

$$\langle p \rangle = \int_{-\infty}^{\infty} dx \psi^*(x, t) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right) \psi(x, t) \quad (2-39)$$

This suggests that the momentum be represented by the differential *operator*

$$p_{op} = \frac{\hbar}{i} \frac{\partial}{\partial x} \quad (2-40)$$

Once we accept this, we can easily calculate

$$\langle p^2 \rangle = \int_{-\infty}^{\infty} dx \psi^*(x, t) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} \right)^2 \psi(x, t) = -\hbar^2 \int_{-\infty}^{\infty} dx \psi^*(x, t) \frac{\partial^2 \psi(x, t)}{\partial x^2}$$

and more generally

$$\langle f(p) \rangle = \int_{-\infty}^{\infty} dx \psi^*(x, t) f\left(\frac{\hbar}{i} \frac{\partial}{\partial x}\right) \psi(x, t) \quad (2-41)$$

The form of the momentum operator in (2-40) raises the question whether the expectation value of the momentum in some states could be imaginary. We can, in fact, show that the expectation value of p is always real. We write

$$\begin{aligned} \langle p \rangle - \langle p \rangle^* &= \int_{-\infty}^{\infty} dx \left[\psi^* \left(\frac{\hbar}{i} \right) \frac{\partial \psi}{\partial x} - \psi \left(-\frac{\hbar}{i} \right) \frac{\partial \psi^*}{\partial x} \right] \\ &= \frac{\hbar}{i} \int_{-\infty}^{\infty} dx \left[\psi^* \frac{\partial \psi}{\partial x} + \psi \frac{\partial \psi^*}{\partial x} \right] = \frac{\hbar}{i} \int_{-\infty}^{\infty} dx \frac{\partial}{\partial x} (\psi^* \psi) = 0 \end{aligned} \quad (2-42)$$

The last step follows from the square integrability of the wave functions. Sometimes one has occasion to use functions that are not square integrable but that obey periodic boundary conditions, such as

$$\psi(x) = \psi(x + L) \quad (2-43)$$

Also, under these circumstances

$$\langle p \rangle - \langle p \rangle^* = \frac{\hbar}{i} \int_0^L dx \frac{\partial}{\partial x} (\psi^* \psi) = \frac{\hbar}{i} (|\psi(L)|^2 - |\psi(0)|^2) = 0$$

An operator whose expectation value for all admissible wave functions is real is called a *hermitian operator*. We see that the momentum operator is hermitian.

The form of the momentum operator allows us to write the Schrödinger equation in the form

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = H\psi(x, t) \quad (2-44)$$

where the operator H , called the *Hamiltonian operator*, is a transcription of the energy

$$H = \frac{p_{op}^2}{2m} + V(x) \quad (2-45)$$

into operator form, using the operator form of p given in (2-38). The Hamiltonian operator is particularly significant in quantum mechanics, and we shall analyze it in detail in many examples.

Wave Function in Momentum Space

Now that we have obtained a way of representing the momentum, we can discuss the physical meaning of $\phi(p)$, which was found to have the form

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi(x, 0) e^{-ipx/\hbar} \quad (2-46)$$

We can calculate

$$\begin{aligned}
 \int_{-\infty}^{\infty} dp \phi^*(p) \phi(p) &= \int_{-\infty}^{\infty} dp \phi^*(p) \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi(x) e^{-ipx/\hbar} \\
 &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi(x) \int_{-\infty}^{\infty} dp \phi^*(p) e^{-ipx/\hbar} \\
 &= \int_{-\infty}^{\infty} dx \psi(x) \psi^*(x) = 1
 \end{aligned} \tag{2-47}$$

This result is known as *Parseval's theorem* in the mathematical literature. It states that if a function is normalized to 1, so is its Fourier transform.

Next consider

$$\begin{aligned}
 \langle p \rangle &= \int_{-\infty}^{\infty} dx \psi^*(x) \frac{\hbar}{i} \frac{\partial \psi(x)}{\partial x} = \int_{-\infty}^{\infty} dx \psi^*(x) \frac{\hbar}{i} \frac{d}{dx} \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \phi(p) e^{ipx/\hbar} \\
 &= \int_{-\infty}^{\infty} dp \phi(p) p \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \psi^*(x) e^{ipx/\hbar} \\
 &= \int_{-\infty}^{\infty} dp \phi^*(p) p \phi(p)
 \end{aligned} \tag{2-48}$$

This result, together with (2-47), strongly suggests that $\phi(p)$ should be interpreted as the wave function in momentum space, with $|\phi(p)|^2$ yielding the probability density for finding the particle with momentum p .

Lest the reader think that in spite of this symmetry between x - and p -space, $p = -i\hbar(\partial/\partial x)$ is an operator and x is not, we note that x is, in fact, an operator too. It happens to have a simple (multiplicative) form in x -space, but if we want to calculate $\langle f(x) \rangle$ in momentum space, then we can show by methods very similar to the ones used earlier that

$$\langle x \rangle = \int_{-\infty}^{\infty} dp \phi^*(p) \left(i\hbar \frac{\partial}{\partial p} \right) \phi(p) \tag{2-49}$$

and more generally

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} dp \phi^*(p) f \left(i\hbar \frac{\partial}{\partial p} \right) \phi(p) \tag{2-50}$$

In other words, the operator x has the representation

$$x = i\hbar \frac{\partial}{\partial p} \tag{2-51}$$

in momentum space.

The following example illustrates some computations for a specific wave function $\psi(x)$.

EXAMPLE 2-4

Consider a particle whose normalized wave function is

$$\begin{aligned}\psi(x) &= 2\alpha\sqrt{\alpha}xe^{-\alpha x} & x > 0 \\ &= 0 & x < 0\end{aligned}$$

- (a) For what value of x does $P(x) = |\psi(x)|^2$ peak?
 (b) Calculate $\langle x \rangle$ and $\langle x^2 \rangle$.
 (c) What is the probability that the particle is found between $x = 0$ and $x = 1/\alpha$?
 (d) Calculate $\phi(p)$ and use this to calculate $\langle p \rangle$ and $\langle p^2 \rangle$.

SOLUTION

- (a) The peak in $P(x)$ occurs when $dP(x)/dx = 0$ —that is, when

$$\frac{d}{dx}(x^2e^{-2\alpha x}) = 2x(1 - \alpha x)e^{-2\alpha x} = 0$$

which is at $x = 1/\alpha$.

- (b)

$$\langle x \rangle = \int_0^\infty dx x(4\alpha^3 x^2 e^{-2\alpha x}) = \frac{1}{4\alpha} \int_0^\infty dy y^3 e^{-y} = \frac{3!}{4\alpha} = \frac{3}{2\alpha}$$

$$\langle x^2 \rangle = \int_0^\infty dx x^2(4\alpha^3 x^2 e^{-2\alpha x}) = \frac{4!}{8\alpha^2} = \frac{3}{\alpha^2}$$

- (c) The desired probability is

$$P = \int_0^{1/\alpha} dx (4\alpha^3)x^2 e^{-2\alpha x} = \frac{1}{2} \int_0^2 dy y^2 e^{-y} = 0.32$$

- (d)

$$\begin{aligned}\phi(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_0^\infty dx e^{-ipx/\hbar} (2\alpha\sqrt{\alpha})xe^{-\alpha x} \\ &= \sqrt{\frac{4\alpha^3}{2\pi\hbar}} \frac{d}{d\alpha} \int_0^\infty dx e^{-(\alpha + ip/\hbar)x} = -\sqrt{\frac{4\alpha^3}{2\pi\hbar}} \frac{1}{(\alpha + ip/\hbar)^2}\end{aligned}$$

From this we can calculate

$$\begin{aligned}\langle p \rangle &= \int_{-\infty}^\infty dp p |\phi(p)|^2 = \frac{4\alpha^3}{2\pi\hbar} \int_{-\infty}^\infty dp \frac{p}{(\alpha^2 + p^2/\hbar^2)^2} = 0 \\ \langle p^2 \rangle &= \frac{4\alpha^3}{2\pi\hbar} \int_{-\infty}^\infty dp \frac{p^2}{(\alpha^2 + p^2/\hbar^2)^2} = \frac{8\alpha^3}{2\pi\hbar} \int_0^\infty dp \frac{p^2}{(\alpha^2 + p^2/\hbar^2)^2}\end{aligned}$$

With the change of variables $p = \hbar\alpha \tan \theta$, we get

$$\langle p^2 \rangle = \frac{4\alpha^2\hbar^2}{\pi} \int_0^{\pi/2} d\theta \sin^2 \theta = \alpha^2\hbar^2$$

The introduction of *operators* brings in a new concern. Products of operators need careful definition, because the order in which they act is important. Consider for example

$$xp\psi(x) = -i\hbar x \frac{d\psi(x)}{dx} \quad (2-52)$$

whereas

$$px\psi(x) = -i\hbar \frac{d}{dx}(x\psi(x)) = -i\hbar x \frac{d\psi(x)}{dx} - i\hbar\psi(x) \quad (2-53)$$

which is different from (2-52). Note that we can deduce

$$[p, x]\psi(x) \equiv (px - xp)\psi(x) = -i\hbar\psi(x) \quad (2-54)$$

Since this is true for all $\psi(x)$, we conclude that we have an *operator relation*, which reads

$$[p, x] = -i\hbar \quad (2-55)$$

This is a *commutation relation*, and it is interesting because it is a relation between operators, independent of what wave function this acts on. The difference between classical physics and quantum mechanics lies in that physical variables are described by operators and these do not necessarily commute. The commutator (2-55) will be seen to lie at the heart of the Heisenberg uncertainty relation.

PROBLEMS

1. Given that $A(k) = N/(k^2 + \alpha^2)$, calculate $\psi(x)$. Plot $A(k)$ and $\psi(x)$ and show that $\Delta k \Delta x > 1$, independent of the choice of α .
2. The relation between the wavelength λ and the frequency ν in a wave guide is given by

$$\lambda = \frac{c}{\sqrt{\nu^2 - \nu_0^2}}$$

What is the group velocity of such waves?

3. For surface tension waves in shallow water, the relation between frequency and wavelength is given by

$$\nu = \sqrt{\frac{2\pi T}{\rho\lambda^3}}$$

where T is the surface tension and ρ is the density. What is the group velocity of the waves?

4. For deep water gravity waves, the relation between frequency and wavelength is given by

$$\nu = \sqrt{\frac{g}{2\pi\lambda}}$$

What is the group velocity of such waves?

5. Consider the problem of spreading of a gaussian wave packet describing a free particle, with

$$\omega = \frac{k^2\hbar}{2m}$$

Calculate the fractional change in the size of the wave packet in one second if

- (a) The packet represents an electron of mass 0.9×10^{-30} kg, with the wave packet having dimensions of 10^{-6} m; 10^{-10} m.
- (b) The packet represents an object of mass 10^{-3} kg and size 0.01 m. [It will be convenient to express the width in units of $\hbar/mc$, where m is the mass of the particle represented by the packet.]

42 Chapter 2 Wave-Particle Duality, Probability, and the Schrödinger Equation

6. A beam of electrons is to be fired over a distance of 10^4 km. If the size of the initial packet is 10^{-3} m, what will be its size upon arrival, if its kinetic energy is (a) 13.6 eV; (b) 100 MeV? [Caution: The relation between kinetic energy (K. E.) and momentum is not always $K. E. = p^2/2m$!]
7. Consider a wave packet for neutrinos, which are massless to a very good approximation, so that $E = pc$. Show that such a wave packet does not spread.
8. Consider a wave function of the form

$$\psi(x) = Ae^{-\mu|x|}$$

Calculate the wave function in momentum space $\phi(p)$.

9. Consider the example in Problem 8. Calculate A so that $\psi(x)$ is properly normalized.
10. Show that the conservation law (2-33) holds when $\psi(x, t)$ obeys eq. (2-23), but only if $V(x)$ is real.
11. Suppose the $V(x)$ is complex. Obtain an expression for

$$\frac{\partial P(x, t)}{\partial t} \quad \text{and} \quad \frac{d}{dt} \int_{-\infty}^{\infty} dx P(x, t)$$

For absorption of particles the last quantity must be negative (since particles disappear, the probability of their being anywhere decreases). What does this tell us about the imaginary part of $V(x)$?

12. Consider the distribution of grades in a class of 60 students, given by

Grades	60	55	50	45	40	35	30	25	20	15	10	5
# students	1	2	7	9	16	13	3	6	2	0	1	0

- (a) Plot a *histogram* of the distribution.
- (b) Calculate the class average.
- (c) Calculate $(\Delta g)^2 = \langle g^2 \rangle - \langle g \rangle^2$.
13. Compare your histogram with a distribution of the form

$$N(g) = Ce^{-(g-\langle g \rangle)^2/(\Delta g)^2}$$

with C chosen that $\sum_g N(g) = 60$.

14. Show that a grade distribution of the form

Grades	60	50	10
# students	7	34	19

leads to the same average grade, but that the dispersion Δg is different. What is it?

15. Consider the wave function obtained in the first example, $2N/(\sin kx)/x$. For what value of N will it be normalized? [Hint: A useful integral is $\int_{-\infty}^{\infty} dt \left(\frac{\sin t}{t} \right)^2 = \pi$.]
16. Consider the wave function

$$\psi(x) = (\alpha/\pi)^{1/4} \exp(-\alpha x^2/2)$$

Calculate $\langle x^n \rangle$ for $n = 1, 2$. Can you quickly write down the result for $\langle x^{17} \rangle$?

17. Calculate $\phi(p)$ for the wave function in problem 16. Calculate $\langle p^n \rangle$ for $n = 1, 2$.

18. Use the definitions $(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2$ and $(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2$ with the results of Problems 16 and 17 to show that $\Delta p \Delta x > \hbar/2$.
19. Make an estimate of the strength of the nuclear potential energy given the following fact: The “size” of the box that roughly describes the nuclear potential is 10^{-15} m, and it takes 8 MeV to eject a particle from this potential well.
 - (a) Use the uncertainty principle to estimate $\langle p^2 \rangle$ for a nucleon in the box, and given the fact that the mass of the nucleon is $M = 1.67 \times 10^{-27}$ kg, estimate the kinetic energy of the nucleon.
 - (b) Since the potential that gives rise to the binding must more than compensate for this, what is the *negative* potential energy?
20. Monochromatic light passes through a shutter that opens for a time $\Delta t = 10^{-10}$ sec. What is the spread in frequencies caused by the shutter?
21. Nuclei typically of size 10^{-14} m frequently emit electrons with energies in the range 1–10 MeV. In the early days of nuclear physics, people believed that electrons “lived” inside the nuclei. Use the uncertainty relation to show that electrons of such energies could not be contained inside the nucleus.
22. Show that eq. (2-49) holds.

Supplement 2-A

The Fourier Integral and Delta Functions

Consider a function $f(x)$ that is periodic, with period $2L$, so that

$$f(x) = f(x + 2L) \quad (2A-1)$$

Such a function can be expanded in a Fourier series in the interval $(-L, L)$, and the series has the form

$$f(x) = \sum_{n=0}^{\infty} A_n \cos \frac{n\pi x}{L} + \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L} \quad (2A-2)$$

We can rewrite the series in the form

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{in\pi x/L} \quad (2A-3)$$

which is certainly possible, since

$$\cos \frac{n\pi x}{L} = \frac{1}{2} (e^{in\pi x/L} + e^{-in\pi x/L})$$

$$\sin \frac{n\pi x}{L} = \frac{1}{2i} (e^{in\pi x/L} - e^{-in\pi x/L})$$

The coefficients can be determined with the help of the orthonormality relation

$$\frac{1}{2L} \int_{-L}^L dx e^{in\pi x/L} e^{-im\pi x/L} = \delta_{mn} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \quad (2A-4)$$

Thus

$$a_n = \frac{1}{2L} \int_{-L}^L dx f(x) e^{-in\pi x/L} \quad (2A-5)$$

Let us now rewrite (2A-3) by introducing Δn , the difference between two successive integers. Since this is unity, we have

$$\begin{aligned} f(x) &= \sum_n a_n e^{in\pi x/L} \Delta n \\ &= \frac{L}{\pi} \sum_n a_n e^{in\pi x/L} \frac{\pi \Delta n}{L} \end{aligned} \quad (2A-6)$$

Let us change the notation by writing

$$\frac{\pi n}{L} = k \quad (2A-7)$$

and

$$\frac{\pi \Delta n}{L} = \Delta k \quad (2A-8)$$

We also write

$$\frac{La_n}{\pi} = \frac{A(k)}{\sqrt{2\pi}} \quad (2A-9)$$

Hence (2A-6) becomes

$$f(x) = \sum \frac{A(k)}{\sqrt{2\pi}} e^{ikx} \Delta k \quad (2A-10)$$

If we now let $L \rightarrow \infty$, then k approaches a continuous variable, since Δk becomes infinitesimally small. If we recall the Riemann definition of an integral, we see that in the limit (2A-10) can be written in the form

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk A(k) e^{ikx} \quad (2A-11)$$

The coefficient $A(k)$ is given by

$$\begin{aligned} A(k) &= \sqrt{2\pi} \frac{L}{\pi} \cdot \frac{1}{2L} \int_{-L}^L dx f(x) e^{-in\pi x/L} \\ &\rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx f(x) e^{-ikx} \end{aligned} \quad (2A-12)$$

Equations (2A-11) and (2A-12) define the Fourier integral transformations. If we insert the second equation into the first we get

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikx} \int_{-\infty}^{\infty} dy f(y) e^{-iky} \quad (2A-13)$$

Suppose now that we interchange, without question, the order of integrations. We then get

$$f(x) = \int_{-\infty}^{\infty} dy f(y) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x-y)} \right] \quad (2A-14)$$

For this to be true, the quantity $\delta(x - y)$ defined by

$$\delta(x - y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x-y)} \quad (2A-15)$$

and called the *Dirac delta function* must be a very peculiar kind of function; it must vanish when $x \neq y$, and it must tend to infinity in an appropriate way when $x - y = 0$, since the range of integration is infinitesimally small. It is therefore not a function of the usual

mathematical sense, but it is rather a “generalized function” or a “distribution.”¹ It does not have any meaning by itself, but it can be defined provided it always appears in the form

$$\int dx f(x) \delta(x - a)$$

with the function $f(x)$ sufficiently smooth in the range of values that the argument of the delta function takes. We will take that for granted and manipulate the delta function by itself, with the understanding that at the end all the relations that we write down only occur under the integral sign.

The following properties of the delta function can be demonstrated:

(i)

$$\delta(ax) = \frac{1}{|a|} \delta(x) \quad (2A-16)$$

This can be seen to follow from

$$f(x) = \int dy f(y) \delta(x - y) \quad (2A-17)$$

If we write $x = a\xi$ and $y = a\eta$, then this reads

$$f(a\xi) = |a| \int d\eta f(a\eta) \delta[a(\xi - \eta)]$$

On the other hand,

$$f(a\xi) = \int d\eta f(a\eta) \delta(\xi - \eta)$$

which implies our result.

(ii) A relation that follows from (2A-16) is

$$\delta(x^2 - a^2) = \frac{1}{2|a|} [\delta(x - a) + \delta(x + a)] \quad (2A-18)$$

This follows from the fact that the argument of the delta function vanishes at $x = a$ and $x = -a$. Thus there are two contributions:

$$\begin{aligned} \delta(x^2 - a^2) &= \delta[(x - a)(x + a)] \\ &= \frac{1}{|x + a|} \delta(x - a) + \frac{1}{|x - a|} \delta(x + a) \\ &= \frac{1}{2|a|} [\delta(x - a) + \delta(x + a)] \end{aligned}$$

More generally, one can show that

$$\delta[f(x)] = \sum_i \frac{\delta(x - x_i)}{|df/dx|_{x=x_i}} \quad (2A-19)$$

where the x_i are the roots of $f(x)$ in the interval of integration.

¹The theory of distributions was developed by the mathematician Laurent Schwartz. An introductory treatment may be found in M. J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions*, Cambridge University Press, Cambridge, England, 1958.

In addition to the representation (2A-15) of the delta function, there are other representations that may prove useful. We discuss several of them.

(a) Consider the form (2A-15), which we write in the form

$$\delta(x) = \frac{1}{2\pi} \lim_{L \rightarrow \infty} \int_{-L}^L dk e^{ikx} \quad (2A-20)$$

The integral can be done, and we get

$$\begin{aligned} \delta(x) &= \lim_{L \rightarrow \infty} \frac{1}{2\pi} \frac{e^{iLx} - e^{-iLx}}{ix} \\ &= \lim_{L \rightarrow \infty} \frac{\sin Lx}{\pi x} \end{aligned} \quad (2A-21)$$

(b) Consider the function $\Delta(x, a)$ defined by

$$\begin{aligned} \Delta(x, a) &= 0 & x < -a \\ &= \frac{1}{2a} & -a < x < a \\ &= 0 & a < x \end{aligned} \quad (2A-22)$$

Then

$$\delta(x) = \lim_{a \rightarrow 0} \Delta(x, a) \quad (2A-23)$$

It is clear that an integral of a product of $\Delta(x, a)$ and a function $f(x)$ that is smooth near the origin will pick out the value at the origin

$$\begin{aligned} \lim_{a \rightarrow 0} \int dx f(x) \Delta(x, a) &= f(0) \lim_{a \rightarrow 0} \int dx \Delta(x, a) \\ &= f(0) \end{aligned}$$

(c) By the same token, any peaked function, normalized to unit area under it, will approach a delta function in the limit that the width of the peak goes to zero. We will leave it to the reader to show that the following are representations of the delta function:

$$\delta(x) = \lim_{a \rightarrow 0} \frac{1}{\pi} \frac{a}{x^2 + a^2} \quad (2A-24)$$

and

$$\delta(x) = \lim_{a \rightarrow \infty} \frac{\alpha}{\sqrt{\pi}} e^{-\alpha^2 x^2} \quad (2A-25)$$

(d) We will have occasion to deal with *orthonormal polynomials*, which we denote by the general symbol $P_n(x)$. These have the property that

$$\int dx P_m(x) P_n(x) w(x) = \delta_{mn} \quad (2A-26)$$

where $w(x)$ may be unity or some simple function, called the weight function. For functions that may be expanded in a series of these orthogonal polynomials, we can write

$$f(x) = \sum_n a_n P_n(x) \quad (2A-27)$$

If we multiply both sides by $w(x)P_m(x)$ and integrate over x , we find that

$$a_m = \int dy w(y)f(y)P_m(y) \quad (2A-28)$$

We can insert this into (2A-27) and, prepared to deal with “generalized functions,” we freely interchange sum and integral. We get

$$\begin{aligned} f(x) &= \sum_n P_n(x) \int dy w(y)f(y)P_n(y) \\ &= \int dy f(y) \left(\sum_n P_n(x)w(y)P_n(y) \right) \end{aligned} \quad (2A-29)$$

Thus we get still another representation of the delta function. Examples of the $P_n(x)$ are Legendre polynomials, Hermite polynomials, and Laguerre polynomials, all of which make their appearance in quantum mechanical problems.

Since the delta function always appears multiplied by a smooth function under an integral sign, we can give meaning to its derivatives. For example,

$$\begin{aligned} \int_{-\varepsilon}^{\varepsilon} dx f(x) \frac{d}{dx} \delta(x) &= \int_{-\varepsilon}^{\varepsilon} dx \frac{d}{dx} [f(x) \delta(x)] - \int_{-\varepsilon}^{\varepsilon} dx \frac{df(x)}{dx} \delta(x) \\ &= - \int_{-\varepsilon}^{\varepsilon} dx \frac{df(x)}{dx} \delta(x) \\ &= - \left(\frac{df}{dx} \right)_{x=0} \end{aligned} \quad (2A-30)$$

and so on. The delta function is an extremely useful tool, and the student will encounter it in every part of mathematical physics.

The integral of a delta function is

$$\begin{aligned} \int_{-\infty}^x dy \delta(y - a) &= 0 \quad x < a \\ &= 1 \quad x > a \\ &\equiv \theta(x - a) \end{aligned} \quad (2A-31)$$

which is the standard notation for this discontinuous function. Conversely, the derivative of the so-called *step function* is the Dirac delta function:

$$\frac{d}{dx} \theta(x - a) = \delta(x - a) \quad (2A-32)$$

Supplement 2-B

A Brief Tutorial on Probability

In this supplement we give a brief discussion of probability. For simplicity we start with discrete events. Consider the toss of a six-faced die. If the die is not perfectly symmetric, the outcome for face n ($n = 1, 2, \dots, 6$) has a probability p_n . In a large number N of tosses, the face “ n ” turns up a_n times, and we say that the probability of getting the face n is

$$p_n = \frac{a_n}{N} \quad (2B-1)$$

Since $\sum a_n = N$, it follows that

$$\sum p_n = 1 \quad (2B-2)$$

Let us now assign to each face a certain “payoff.” We assign points in the following manner: 1 point when face 1 turns up, 2 points when face 2 turns up, and so on. In N tosses we get n points a_n times, so that the total number of points is $\sum na_n$. This, of course, grows with N . We thus focus on the average value (points per toss), so that

$$\langle n \rangle = \frac{1}{N} \sum na_n = \sum np_n \quad (2B-3)$$

We may be interested in an average of n^2 , say, and in that case we calculate

$$\langle n^2 \rangle = \frac{1}{N} \sum n^2 a_n = \sum n^2 p_n \quad (2B-4)$$

and so on.

When we do not have a discrete outcome, we must deal with densities. To be specific, consider a quantity that varies continuously, for example the height of a population of students. We can make this discrete by making a *histogram* (Fig. 2B-1) plotting the height by listing people’s heights in 10-cm intervals. Nobody will be exactly 180 cm or 190 cm tall, so we just group people, and somehow round things up at the edges. We may want a finer detail, and list the heights in 1-cm intervals or 1-mm intervals, which will continue to make things discrete, but as the intervals become smaller, the histogram resembles more and more a continuous curve. Let us take some interval, dx , and treat it as infinitesimal. The number of people whose height lies between x and $x + dx$ is $n(x) dx$. The proportionality to dx is obvious: twice as many people will fall into the interval $2dx$ as fall into dx . It is here that the infinitesimal character of dx comes in: we do not need to decide whether $n(x) dx$ or $n(x + dx/2) dx$ is to be taken, since we treat $(dx)^2$ as vanishingly small. If the total population has N members, we can speak of the probability of falling into the particular interval as

$$\frac{1}{N} n(x) dx = p(x) dx \quad (2B-5)$$

W-12 Supplement 2-B A Brief Tutorial on Probability

$p(x)$ is called a *probability density*. Since the total number of students is N , we have

$$\int dx n(x) = N \quad (2B-6)$$

or

$$\int dx p(x) = 1 \quad (2B-7)$$

If we want the probability of finding a student of height between a (say, 150 cm) and b (say, 180 cm), we calculate

$$P(a \leq x \leq b) = \int_a^b dx p(x) \quad (2B-8)$$

The average height is calculated in the same way as for the discrete case

$$\langle x \rangle = \int xp(x) dx \quad (2B-9)$$

and we can calculate other quantities such as

$$\langle x^2 \rangle = \int x^2 p(x) dx \quad (2B-10)$$

and so on. Instead of calling these quantities averages, we call them *expectation values*, a terminology that goes back to the roots of probability theory in gambling theory.

We are often interested in a quantity that gives a measure of how the heights, say, are distributed about the average. The deviations from the average must add up to zero. Formally it is clear that

$$\langle x - \langle x \rangle \rangle = 0 \quad (2B-11)$$

since the average of any number, including $\langle x \rangle$, is just that number. We can, however, calculate the average value of the square of the deviation, and this quantity will not be zero. In fact,

$$\begin{aligned} \langle (x - \langle x \rangle)^2 \rangle &= \langle x^2 - 2x\langle x \rangle + \langle x \rangle^2 \rangle = \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2 \\ &= \langle x^2 \rangle - \langle x \rangle^2 = (\Delta x)^2 \end{aligned} \quad (2B-12)$$

This quantity is called the *dispersion*.

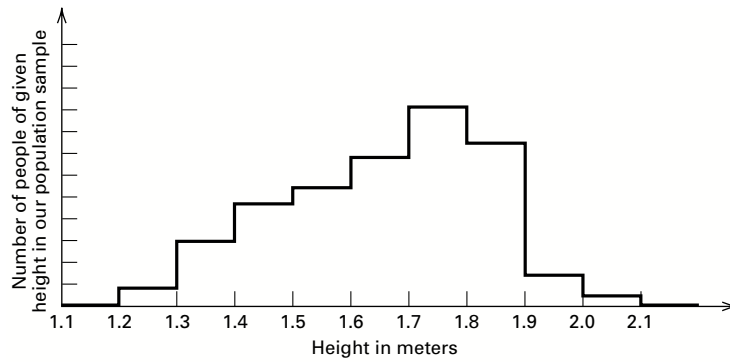


Figure 2B-1
Example of a histogram.

The dispersion is used in quantum mechanics to define the *uncertainty*, so that in the Heisenberg uncertainty relations proper definitions of $(\Delta x)^2$ and $(\Delta p)^2$ are given by

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 \quad (2B-13)$$

and

$$(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2 \quad (2B-14)$$

There is one more point to be made that is directly relevant to quantum physics. We may illustrate it by going back to the population of students. Let us ask for the probability that a randomly chosen student's height lies between 160 cm and 170 cm, *and* that his or her telephone number ends in an even integer. Assuming that there is nothing perverse about how telephone numbers are distributed in the community, half the students will fall into the even category, and thus the desired probability is $P(160 \leq x \leq 170) \times (1/2)$. This is just an example of a general rule *that the probability of two (or more) uncorrelated events is the product of the individual probabilities*.

Thus, if the probability that a particle is in some state “ n ” in our laboratory is $P(n)$, and the probability that a different particle in a laboratory across the country is in a state “ m ” is $P(m)$, then the joint probability that one finds “ n ” in the first laboratory *and* “ m ” in the second laboratory is

$$P(m, n) = P(n)P(m) \quad (2B-15)$$

We will find that a similar result holds for probability amplitudes—that is, for wave functions.

Chapter 3

Eigenvalues, Eigenfunctions, and the Expansion Postulate

Our discussion in Chapter 2 led us to the Schrödinger equation for the wave function $\psi(x, t)$ and its interpretation. In this chapter we analyze the equation in some detail. We find that for potentials $V(x)$ that do not depend on time, the equation reduces to a simple equation, the *time-independent* Schrödinger equation. We solve this equation for some very simple, special cases and obtain some insights into the meaning of the solutions, which we then extend to the general case of an arbitrary potential $V(x)$.

3-1 THE TIME-INDEPENDENT SCHRÖDINGER EQUATION

Let us consider the time-dependent Schrödinger equation obtained in Chapter 2,

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x)\psi(x, t) \equiv H\psi(x, t) \quad (3-1)$$

We introduced a symbol H for the operator on the right-hand side of this equation, which stands for *Hamiltonian*. Equation (3-1) is a partial differential equation. Such equations are much more difficult to deal with than ordinary differential equations in one variable. In our case, because $V(x)$ does not depend on the variable t , it is possible to reduce the partial differential equation to two ordinary differential equations. We write

$$\psi(x, t) = T(t)u(x) \quad (3-2)$$

which implies that

$$i\hbar u(x) \frac{dT(t)}{dt} = \left\{ -\frac{\hbar^2}{2m} \frac{d^2 u(x)}{dx^2} + V(x)u(x) \right\} T(t)$$

Division by $u(x)T(t)$ gives us

$$i\hbar \frac{1}{T(t)} \frac{dT(t)}{dt} = \frac{1}{u(x)} \left\{ -\frac{\hbar^2}{2m} \frac{d^2 u(x)}{dx^2} + V(x)u(x) \right\} \quad (3-3)$$

Since the two sides of the equation depend on different variables, this can only be satisfied if both sides are constant. We denote the constant by E . The solution of

$$i\hbar \frac{dT(t)}{dt} = ET(t) \quad (3-4)$$

is

$$T(t) = Ce^{-iEt/\hbar} \quad (3-5)$$

where C is a constant. The other equation is

$$-\frac{\hbar^2}{2m} \frac{d^2 u(x)}{dx^2} + V(x)u(x) = Eu(x) \quad (3-6)$$

This equation is frequently called the *time-independent Schrödinger equation*.

In classical mechanics the motion of a particle is determined by the force $F(x) = -dV(x)/dx$, so that adding a constant V_0 to the potential has no effect. We can see that the same holds in quantum mechanics. If $V(x)$ is replaced by $V(x) + V_0$, then all we need to do is to replace the label E in (3-4) by $E + V_0$. This means that

$$T(t) = Ce^{-i(E+V_0)t/\hbar}$$

and

$$E + V_0 = \frac{1}{u(x)} \left\{ \frac{\hbar^2}{2m} \frac{d^2 u(x)}{dx^2} + (V(x) + V_0)u(x) \right\}$$

so that (3-6) is unchanged. The presence of V_0 merely changes the solution $\psi(x, t)$ by a phase factor $e^{-iV_0 t/\hbar}$.

Let us return to (3-6). Its character is really different from that of (3-1). Equation (3-1) describes the time evolution of the wave function $\psi(x, t)$, whereas (3-6) is an *eigenvalue equation*. What do we mean by that?

3-2 EIGENVALUE EQUATIONS

A discussion of eigenvalue equations needs a more careful consideration of operators than was presented in Chapter 2. Most generally, an operator acting on a function maps it into another function. Let us consider some examples:

$$\begin{aligned} Of(x) &= f(x) + x^2 \\ Of(x) &= [f(x)]^2 \\ Of(x) &= f(3x^2 + 1) \\ Of(x) &= [df(x)/dx]^3 \\ Of(x) &= df(x)/dx - 2f(x) \\ Of(x) &= \lambda f(x) \end{aligned} \quad (3-7)$$

All of these examples share the property that given a function $f(x)$, there is a rule that determines $Of(x)$ for us. There is a special class of operators, called *linear operators* (we denote these operators by L to distinguish them from the general operators O). These have the property that

$$L[f_1(x) + f_2(x)] = Lf_1(x) + Lf_2(x) \quad (3-8)$$

and, with c an arbitrary complex number,

$$Lcf(x) = cLf(x) \quad (3-9)$$

Thus, in our list, the third, fifth, and sixth operators are linear.

A linear operator will map one function into another, as in the example

$$Lf(x) = \frac{df(x)}{dx} - 2f(x)$$

It is instructive to think of the functions as analogous to vectors in a three-dimensional space. Here, the action of an operator is to transform a vector into another vector. In the special case that the vectors are all of unit length, an operator will transform one point on a unit sphere into another.

Let us consider a special example. We define an operator that acts on a unit vector in the following way: It rotates the vector about the polar axis in such a way that the tip of the vector moves a distance π on a circle of fixed latitude (see Fig. 3-1). If we call this operator R , then the eigenvalue equation will be

$$\mathbf{r}' = R\mathbf{r} = \lambda\mathbf{r} \quad (3-10)$$

Since the length of the vector does not change, $r'^2 = r^2 = 1$ so that $\lambda^2 = 1$ and thus $\lambda = \pm 1$. We can now study the “eigenfunctions”—that is, the directions of the vectors $\mathbf{r}$ that satisfy the above equation. For $\lambda = 1$, the vector will return to its starting point if the distance π is equal to an integer multiple of the circumference of the fixed latitude circle. Since for a sphere of unit radius the radius of the circle is $\sin \theta$, and so its circumference is $2\pi \sin \theta$, we must restrict the vector to values of θ such that

$$\pi = n(2\pi \sin \theta)$$

that is,

$$\sin \theta = 1/2n \quad n = 1, 2, 3 \dots$$

Thus the eigenvectors lie on a series of circles starting at 30° and approaching the north pole, and a similar series of circles starting at 150° and approaching the south pole. There is only one allowed value of θ for the eigenvalue $\lambda = -1$. A change of sign can only happen for a vector lying on the equator, so that $\theta = \pi/2$.

Equation (3-6) may be written in the form

$$Hu_E(x) = Eu_E(x) \quad (3-11)$$

This equation states that H , the Hamiltonian operator acting on a special class of functions, will give back the function that it is acting on, multiplied by a constant, called the *eigenvalue*. The solution of the equation depends on E and we have therefore labeled it with an E . The solution $u_E(x)$ is called the *eigenfunction* corresponding to the eigenvalue E of the operator H . We shall see that eigenvalues can form a continuum or be discrete. Because the operator H has the form

$$H = \frac{p_{op}^2}{2m} + V(x)$$

the eigenvalues are called *energy eigenvalues* and the eigenfunctions are *energy eigenfunctions*.

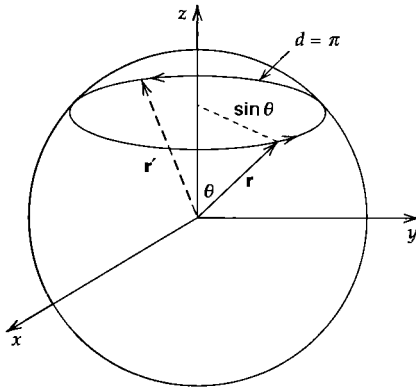


Figure 3-1 The action of R on a unit vector on a sphere.

EXAMPLE 3-1

Solve the eigenvalue problem

$$i \frac{df(\varphi)}{d\varphi} = \lambda f(\varphi)$$

subject to the condition

$$f(\varphi) = f(\varphi + 2\pi)$$

SOLUTION The eigenvalue equation may be written in the form

$$i \frac{1}{f(\varphi)} \frac{df(\varphi)}{d\varphi} = \lambda$$

which is equivalent to

$$\frac{d}{d\varphi} \ln f(\varphi) = -i\lambda$$

Integration gives

$$\ln f(\varphi) = -i\lambda\varphi + \text{constant}$$

that is,

$$f(\varphi) = f(0)e^{-i\lambda\varphi}$$

The periodic condition implies that

$$e^{-i\lambda(2\pi)} = 1$$

This means that the allowed eigenvalues are $\lambda = 0, \pm 1, \pm 2, \pm 3, \dots$

The solution of the Schrödinger equation that we obtained is of the form $u_E(x)e^{-iEt\hbar}$, but since (3-1) is a linear equation, a sum of the solution of the above form is also a solution. Permissible values of E may be discrete, E_n ($n = 1, 2, 3, \dots$) or continuous. If we describe the eigenfunctions corresponding to the discrete eigenvalues E_n by $u_n(x)$ and those corresponding to continuous eigenvalues E by $u_E(x)$, then the most general solution of (3-1) is

$$\psi(x, t) = \sum_n C_n u_n(x) e^{-iE_n t \hbar} + \int dE C(E) u_E(x) e^{-iEt \hbar} \quad (3-12)$$

Here the C_n are arbitrary constants and the $C(E)$ are arbitrary functions of E , only constrained by the condition that $\psi(x, t)$ be square-integrable.

The separation of (3-1) would not have worked if the potential energy V depended on the time. We will see later that when V does depend on t , then energy is not conserved, so that it is no longer a constant of the motion. We now turn to the discussion of a simple but instructive example.

3-3 THE EIGENVALUE PROBLEM FOR A PARTICLE IN A BOX

We consider (3-6) with

$$\begin{aligned} V(x) &= \infty & x < 0 \\ &= 0 & 0 < x < a \\ &= \infty & a < x \end{aligned} \quad (3-13)$$

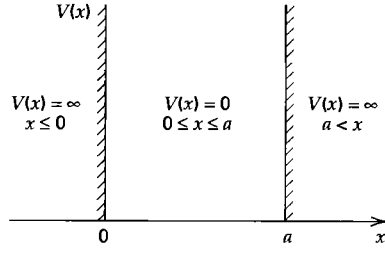


Figure 3-2 The infinite well.

(See Fig. 3-2.) We see from (3-6) that

$$\begin{aligned} u(x) &= 0 & x < 0 \\ &= 0 & a < x \end{aligned} \quad (3-14)$$

and inside the box, where $V(x) = 0$, we rewrite (3-6) in the form

$$\frac{d^2 u(x)}{dx^2} + \frac{2mE}{\hbar^2} u(x) = 0 \quad (3-15)$$

First we notice that if $E < 0$, then the equation takes the form

$$\frac{d^2 u(x)}{dx^2} - \kappa^2 u(x) = 0$$

where $\kappa^2 = 2m|E|/\hbar^2$. The most general solution is some linear combination of $e^{\kappa x}$ and $e^{-\kappa x}$, and although $\sinh \kappa x$ vanishes at $x = 0$, it does not vanish at $x = a$. Thus we cannot have $E < 0$. With positive E , and the notation

$$k^2 = \frac{2mE}{\hbar^2} \quad (3-16)$$

(3-15) takes the form

$$\frac{d^2 u(x)}{dx^2} + k^2 u(x) = 0 \quad (3-17)$$

The most general solution is of the form $A \sin kx + B \cos kx$, but the requirement that $u(0) = 0$ limits us to

$$u(x) = A \sin kx \quad (3-18)$$

The condition $u(a) = 0$ implies that

$$ka = n\pi \quad n = 1, 2, 3, \dots \quad (3-19)$$

Thus the energy eigenvalues are

$$E_n = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2ma^2} \quad n = 1, 2, 3, \dots \quad (3-20)$$

It is easy to check that the solutions are normalized if $A = \sqrt{2/a}$, so that

$$u_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a} \quad (3-21)$$

The solutions have the property that

$$\begin{aligned}
 \int_0^a dx u_n^*(x) u_m(x) &= \int_0^a dx \frac{2}{a} \sin \frac{n\pi x}{a} \sin \frac{m\pi x}{a} \\
 &= \frac{1}{a} \int_0^a dx \left\{ \cos \frac{(n-m)\pi x}{a} - \cos \frac{(n+m)\pi x}{a} \right\} \\
 &= \frac{\sin(n-m)\pi}{(n-m)\pi} - \frac{\sin(n+m)\pi}{(n+m)\pi} \\
 &= 0 \quad \text{when } n \neq m \\
 &= 1 \quad \text{when } n = m
 \end{aligned} \tag{3-22}$$

The conclusion that

$$\int_0^a dx u_n^*(x) u_m(x) = \delta_{mn} \tag{3-23}$$

implies that eigenfunctions corresponding to different eigenvalues are *orthogonal*. If the eigenfunctions are properly normalized, as they are here, (3-23) is called the *orthonormality condition*. Since the solutions are real, the complex conjugation in that equation is not really necessary, but is inserted for consistency with the more general statement that applies when the eigenfunctions are complex. This will be established in Chapter 5.

We can extract some physical information from the eigensolutions:

1. The state of lowest energy, the *ground state*, is described by $u_1(x)$, and the lowest energy is

$$E_1 = \frac{\pi^2 \hbar^2}{2ma^2} \tag{3-24}$$

Note that classically the lowest energy would be that of a particle at rest in the hole, and for $p = 0$ and $V(x) = 0$, the sum of the kinetic and potential energies would be zero. Here we see the presence of a minimum energy.

2. Since the solutions are real,

$$\begin{aligned}
 \langle p \rangle &= \int_0^a dx u_n(x) \left(-i\hbar \frac{du_n(x)}{dx} \right) = -i\hbar \int_0^a dx \frac{d}{dx} \frac{(u_n(x))^2}{2} \\
 &= -\frac{i\hbar}{2} [u_n^2(a) - u_n^2(0)] = 0
 \end{aligned} \tag{3-25}$$

On the other hand, $\langle p^2 \rangle$ does not vanish. In fact, inside the box $p^2 = 2mE$, so that

$$\langle p^2 \rangle = 2mE_n = \frac{\hbar^2 \pi^2 n^2}{a^2} \tag{3-26}$$

for the eigenfunctions $u_n(x)$.

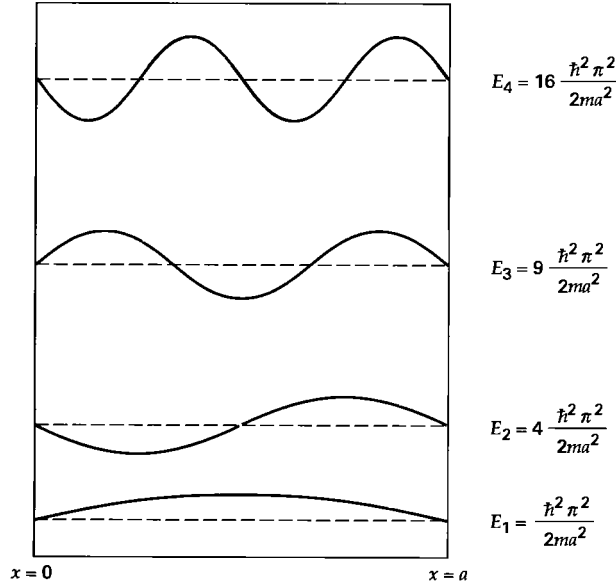


Figure 3-3 Eigensolutions for particles in a box.

3. The larger the number of nodes in a solution, the higher is its kinetic energy, as can be seen in Fig. 3-3. This is understandable since the kinetic energy grows with the curvature of the solutions. We see this by looking at the general form of the expectation value of the kinetic energy. It is

$$\begin{aligned}
 \langle K \rangle &= \frac{\langle p^2 \rangle}{2m} = -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx u^*(x) \frac{d^2 u(x)}{dx^2} \\
 &= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \left\{ \frac{d}{dx} \left(u^*(x) \frac{du(x)}{dx} \right) - \frac{du^*(x)}{dx} \frac{du(x)}{dx} \right\} \quad (3-27) \\
 &= \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \left| \frac{du(x)}{dx} \right|^2
 \end{aligned}$$

The first term in the middle line vanishes because $u(x)$ and its derivative vanish at infinity. We see that $\langle K \rangle$ is large if $u(x)$ oscillates a lot.

4. The eigenfunctions shown in Fig. 3-3 exhibit a pattern. Those corresponding to odd values of n are unaltered if they are reflected in a mirror placed along the mid-point of the well, while those corresponding to even values of n change sign on reflection. It is therefore instructive to exhibit this symmetry explicitly. We do so by displacing the well so that the “mirror” lies along the vertical axis at $x = 0$. This puts the boundaries of the box at $x = -a/2$ and $x = +a/2$. This shift is accomplished by the change of variables

$$x \rightarrow x - \frac{a}{2}$$

We then get

$$\sin \frac{n\pi x}{a} \rightarrow \sin \left(\frac{n\pi x}{a} - \frac{n\pi}{2} \right) = \sin \frac{n\pi x}{a} \cos \frac{n\pi}{2} - \cos \frac{n\pi x}{a} \sin \frac{n\pi}{2} \quad (3-28)$$

For $n = 1, 3, 5, \dots$, the solutions (aside from a sign in front, absorbed in the phase of the normalization constant) are proportional to $\cos n\pi x/a$, which are *even* functions of x . For $n = 2, 4, 6, \dots$, the solutions are proportional to $\sin n\pi x/a$, which are *odd* functions of x , and which therefore vanish on the symmetry axis, now labeled with $x = 0$.

By making the shift we have obtained the eigensolutions for the infinite well bounded by $x = \pm a/2$. They are

$$\sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a} \quad \text{with } n = 2, 4, 6, \dots$$

and

$$\sqrt{\frac{2}{a}} \cos \frac{n\pi x}{a} \quad \text{with } n = 1, 3, 5, \dots$$

EXAMPLE 3-2

Consider a one-dimensional box of width 1 mm. Roughly, what value of n corresponds to a state of energy 0.01 eV?

SOLUTION We calculate

$$\begin{aligned} n^2 &= \frac{2ma^2E}{\hbar^2\pi^2} = \\ &= \frac{(2 \times 0.9 \times 10^{-30} \text{ kg})(10^{-3} \text{ m})^2(10^{-2} \text{ eV})(1.6 \times 10^{-19} \text{ J/eV})}{(1.05 \times 10^{-34} \text{ J}\cdot\text{s})^2(3.14)^2} \\ &= 2.65 \times 10^8 \end{aligned}$$

The value of n is roughly

$$n = 1.63 \times 10^4$$

EXAMPLE 3-3

Calculate the density of states for Example 3-2 in the vicinity of 0.01 eV. The density of states is defined by dn/dE . What is the number of states within an interval of 10^{-4} eV about the energy of 10^{-2} eV?

SOLUTION We have

$$E = (\text{const})n^2$$

so that

$$\ln E = \ln(\text{const}) + 2 \ln n$$

From this we get

$$\frac{dE}{E} = 2 \frac{dn}{n}$$

so that

$$\frac{dn}{dE} = \frac{n}{2E}$$

Numerically we get

$$\frac{dn}{dE} = \frac{1.63 \times 10^4}{2 \times 10^{-2} \text{ eV}} = 0.82 \times 10^6 \text{ states/eV}$$

The number of states in a 10^{-4} eV interval is therefore 82.

3-4 THE EXPANSION POSTULATE AND ITS PHYSICAL INTERPRETATION

Fourier's theorem states that any function $\psi(x)$ that satisfies the boundary conditions $\psi(0) = \psi(a) = 0$ can be written in the form

$$\psi(x) = \sum_{n=1}^{\infty} C_n \sin \frac{n\pi x}{a} \quad (3-29)$$

Since the eigenfunctions of H for the infinite well are proportional to $\sin n\pi x/a$, we write the preceding expression in terms of the eigenfunctions $u_n(x)$:

$$\psi(x) = \sum_{n=1}^{\infty} A_n u_n(x) \quad (3-30)$$

The orthonormality relation (3-23) can be used to find A_n . We see that

$$\begin{aligned} \int_0^a dx u_m^*(x) \psi(x) &= \int_0^a dx u_m^*(x) \sum_{n=1}^{\infty} A_n u_n(x) \\ &= \sum_{n=1}^{\infty} A_n \int_0^a dx u_m^*(x) u_n(x) = \sum_{n=1}^{\infty} A_n \delta_{mn} = A_m \end{aligned}$$

so that

$$A_m = \int_0^a dx u_m^*(x) \psi(x) \quad (3-31)$$

As in our discussion of the free wave packet, we can calculate the time development of this arbitrary initial wave function $\psi(x)$. Since each of the eigenfunctions $u_n(x)$ acquires its own time dependence $e^{-iE_n t/\hbar}$, we have quite generally

$$\psi(x, t) = \sum A_n u_n(x) e^{-iE_n t/\hbar} \quad (3-32)$$

The expansion has its counterpart in vector analysis. Consider an n -dimensional space. The set of orthogonal unit vector $\mathbf{i}_m$ plays the role of a complete set of eigenfunctions, *complete* in the sense that any vector $\mathbf{a}$ in this space may be expanded in terms of these unit vectors

$$\mathbf{a} = \sum_{m=1}^n a_m \mathbf{i}_m$$

Since the unit vectors form an orthogonal basis, we have

$$\mathbf{i}_m \cdot \mathbf{i}_k = \delta_{mk}$$

The coefficients in the expansion a_k may be obtained by taking a dot product with unit vectors

$$a_k = \mathbf{i}_k \cdot \mathbf{a}$$

The analogy has its peculiarity: The dot product is replaced by an integral over the variable x . In spite of this, the analogy between wave functions and vectors, between eigenfunctions and a basis of unit vectors will turn out to be very instructive.

Interpretation of the Expansion Coefficients

To interpret the coefficients A_n we calculate the expectation value of the energy in an arbitrary state $\psi(x)$. Since inside the box $H = p^2/2m$, and outside the box $\psi(x) = 0$, and since

$$Hu_n(x) = E_n u_n(x)$$

It follows that

$$\begin{aligned} \langle H \rangle &= \int_0^a dx \psi^*(x) H \psi(x) = \int_0^a dx \psi^*(x) H \sum_n A_n u_n(x) \\ &= \sum_n A_n \int_0^a dx \psi^*(x) E_n u_n(x) \\ &= \sum_n E_n |A_n|^2 \end{aligned} \quad (3-33)$$

We also note that

$$\int_0^a dx \psi^*(x) \psi(x) = 1 \quad (3-34)$$

implies that

$$1 = \int_0^a dx \psi^*(x) \sum_n A_n u_n(x) = \sum_n A_n A_n^* = \sum_n |A_n|^2 \quad (3-35)$$

To interpret $|A_n|^2$, we note that an energy measurement can only yield one of the eigenvalues. This statement was implicit in the starting point of Bohr's description of the stationary states of the atom. We shall take it to be a postulate of quantum mechanics that a measurement of the energy must be one of the eigenvalues of the energy operator. Under these circumstances, a measurement of the average energy of a system will give us

$$E_{av} = \langle H \rangle = \sum_n E_n p_n$$

where p_n is the probability that an individual measurement yields E_n . The probabilities must add up to one, so that

$$\sum_n p_n = 1$$

Comparison with (3-33) and (3-35) immediately suggests that $|A_n|^2$, where

$$A_n = \int_0^a dx u_n^*(x) \psi(x) \quad (3-36)$$

be interpreted as *the probability that a measurement of the energy for the state $\psi(x)$ yields the eigenvalue E_n* .

The assertion that any single energy measurement must yield an eigenvalue E_n has further implications. Suppose such a measurement is made, and E_n results. A repetition of the measurement must yield the same result. (Otherwise, how else could we check that the measurement was carried out correctly?) This, however, means that after the first measurement the system *had to be* in the eigenstate $u_n(x)$, since that is the only way to ensure that the second measurement will give E_n with certainty. This implies that a *measurement projects the initial state into an eigenstate of the observable (the energy here) that is being measured*.

This assertion is not peculiar to the problem of a particle in a box. It holds for more general systems in which there is a potential energy $V(x)$, and also for observables other than the energy, such as the momentum, angular momentum, and so on.

EXAMPLE 3-4

Consider the eigenvalue problem discussed in Example 3-1. The eigenvalues are $n = 0, \pm 1, \pm 2, \dots$ and the normalized eigenfunctions are

$$u_n(\varphi) = \frac{1}{\sqrt{2\pi}} e^{in\varphi}$$

with the normalization determined by the condition that

$$\int_0^{2\pi} d\varphi |u_n(\varphi)|^2 = 1$$

We may view φ as the coordinate of a particle moving on a unit circle. Suppose the normalized wave function of the particle is

$$\psi(\varphi) = N \cos^2 \varphi$$

What is the probability that a measurement of the eigenvalue yields $n = -2$?

SOLUTION We need to expand the wave function ψ in terms of the normalized eigenfunction. We do this by a direct expansion rather than by using the orthonormality of the eigenfunctions, since this allows us to look at other possibilities.

$$\begin{aligned} \psi(\varphi) &= N \left(\frac{e^{i\varphi} + e^{-i\varphi}}{2} \right)^2 = \frac{N}{4} (e^{2i\varphi} + 2 + e^{-2i\varphi}) \\ &= \frac{N}{4} \sqrt{2\pi} (u_2 + 2u_0 + u_{-2}) \end{aligned}$$

Thus the coefficient of u_{-2} is $\sqrt{(2\pi)N}/4$, and the probability is $P(-2) = \pi N^2/8$. We need only to find the value of N^2 . It is determined by

$$N^2 \int_0^{2\pi} d\varphi \cos^4 \varphi = 1$$

One can do the integral, but another way of doing it in simple cases is to note that $P(2) = P(-2) = \pi N^2/8$ and $P(0) = 4\pi N^2/8$. Since all of these add up to 1, we get

$$2 \frac{\pi N^2}{8} + \frac{4\pi N^2}{8} = 1$$

so that $N^2 = 4/3\pi$, and therefore

$$P(-2) = \frac{1}{6}$$

EXAMPLE 3-5

Consider a particle in a box. Its wave function is given by

$$\begin{aligned}\psi(x) &= A(x/a) & 0 < x < a/2 \\ &= A(1 - x/a) & a/2 < x < a\end{aligned}$$

where $A = \sqrt{12/a}$ so as to satisfy $\int_0^a dx |\psi(x)|^2 = 1$. Calculate the probability that a measurement of the energy yields the eigenvalue E_n .

SOLUTION We want to calculate A_n in the expansion

$$\begin{aligned}A_n &= \int_0^a dx \psi(x) \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a} \\ &= \frac{\sqrt{24}}{a} \left[\int_0^{a/2} dx \left(\frac{x}{a} \right) \sin \frac{n\pi x}{a} + \int_{a/2}^a dx \left(1 - \frac{x}{a} \right) \sin \frac{n\pi x}{a} \right]\end{aligned}$$

With the change of variables $\pi x/a = u$ in the first integral and $\pi x/a = \pi - u$ in the second integral, we get

$$A_n = \frac{\sqrt{24}}{\pi} \int_0^{\pi/2} du \frac{u}{\pi} \sin nu (1 - (-1)^n)$$

The A_n for n even vanish because of the last factor. The integral is easily calculated, and we get, for n odd only,

$$A_n = \frac{\sqrt{24}}{\pi} 2 \frac{1}{\pi n^2} (-1)^{n+1}$$

so that

$$\begin{aligned}|A_n|^2 &= \frac{96}{\pi^4 n^4} & \text{for } n \text{ odd} \\ &= 0 & \text{for } n \text{ even}\end{aligned}$$

One can easily check, using the fact that $\sum_{\text{all}} n^{-4} = \pi^4/90$ and

$$\sum_{\text{all}} n^{-4} = \sum_{\text{even}} n^{-4} + \sum_{\text{odd}} n^{-4} = \sum_{\text{odd}} n^{-4} + (1/16) \sum_{\text{all}} n^{-4}$$

that the sum of all the probabilities is 1:

$$\frac{96}{\pi^4} \sum_{\text{odd}} n^{-4} = \frac{96}{\pi^4} \left(1 - \frac{1}{16} \right) \sum_{\text{all}} n^{-4} = \frac{96}{\pi^4} \cdot \frac{15}{16} \cdot \frac{\pi^4}{90} = 1$$

3-5 MOMENTUM EIGENFUNCTION AND THE FREE PARTICLE

The energy operator H is not the only one that has eigenfunctions and eigenvalues. Let us solve the eigenvalue equation for the *momentum operator*

$$p_{\text{op}} u_p(x) = p u_p(x) \quad (3-37)$$

Since $p_{\text{op}} = (\hbar/i)(d/dx)$, this reads

$$\frac{du_p(x)}{dx} = \frac{ip}{\hbar} u_p(x) \quad (3-38)$$

The solution to this equation is

$$u_p(x) = C e^{ipx/\hbar} \quad (3-39)$$

with C a constant to be determined by normalization, and the eigenvalue p real, so that the eigenfunction does not blow up at either $+\infty$ or $-\infty$. This is the only constraint on p : we say that p_{op} has a *continuous spectrum*. We might, by analogy with (3-23), expect that the eigenfunctions obey orthonormality conditions. We see that

$$\begin{aligned} \int_{-\infty}^{\infty} dx u_p^*(x) u_{p'}(x) &= |C|^2 \int_{-\infty}^{\infty} dx e^{i(p-p')x/\hbar} \\ &= 2\pi |C|^2 \hbar \delta(p - p') \end{aligned} \quad (3-40)$$

With the choice

$$u_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \quad (3-41)$$

(3-40) reads

$$\int_{-\infty}^{\infty} dx u_p^*(x) u_{p'}(x) = \delta(p - p') \quad (3-42)$$

This differs from (3-23) only in that the Kronecker δ_{mn} , appropriate for discrete indices is replaced by a Dirac delta function $\delta(p - p')$ for the continuous indices.

The statement that any wave packet $\psi(x)$ may be expanded in terms of a complete set of eigenfunctions can also be established here. The analog of (3-30) must take into account that we are summing over a continuous index p , so that we write

$$\psi(x) = \int_{-\infty}^{\infty} dp \phi(p) \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} \quad (3-43)$$

According to the interpretation implicit in (3-36), $|\phi(p)|^2$, where

$$\phi(p) = \int_{-\infty}^{\infty} dx \left(\frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} \right)^* \psi(x) \quad (3-44)$$

gives the probability that a measurement of the momentum for an arbitrary packet $\psi(x)$ yields the eigenvalue p . In this way we justify the conjecture made about $\phi(p)$ in Chapter 2, that $\phi(p)$ is the wave function in momentum space.

EXAMPLE 3-6

A particle is in the ground state of a box with sides at $x = 0$ and $x = a$. Suddenly the walls of the box are moved to $\pm\infty$, so that the particle is free. What is the probability that the particle has momentum in the range $(p, p + dp)$? Since a free particle of momentum p has energy $p^2/2m$, which need not equal the ground-state energy, energy is not conserved. How is this possible?

SOLUTION Our initial wave function has the form

$$\psi(x) = \sqrt{\frac{2}{a}} \sin \frac{\pi x}{a} \quad 0 \leq x \leq a$$

As we saw in (3-44) the probability amplitude that the particle in this state has momentum p is given by

$$\phi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_0^a dx e^{-ipx/\hbar} \sqrt{\frac{2}{a}} \sin \frac{\pi x}{a}$$

The limits on the integrals come from the condition that $\psi(x)$ vanishes to the left of $x = 0$ and to the right of $x = a$, since that is where the walls initially stood. The integral is easy to evaluate. We get

$$\phi(p) = -e^{-pa/2\hbar} \frac{2\pi/a}{\sqrt{\pi\hbar a} (\pi/a)^2 - (p/\hbar)^2} \cos pa/2\hbar$$

so that

$$|\phi(p)|^2 dp = \frac{4\pi}{\hbar a^3} \frac{\cos^2(pa/2\hbar)}{((\pi/a)^2 - (p/\hbar)^2)^2} dp$$

Energy is not conserved because there is actually a time dependence in the potential energy. In fact, $V(x)$ changes very rapidly from the walls at $x = 0$ and $x = a$, to walls at $\pm\infty$.

Note that (1) the probability drops very rapidly once p gets to be much larger than $\hbar\pi/a$, and (2) the probability density does not become infinite at $p = \hbar\pi/a$ because the numerator vanishes at that point.

Normalization of the Free Particle Wave Function

Let us now turn to the free particle Hamiltonian. When $V(x)$ is zero everywhere, the energy eigenvalue equation reads

$$\frac{d^2 u(x)}{dx^2} + \frac{2mE}{\hbar^2} u(x) = \frac{d^2 u(x)}{dx^2} + k^2 u(x) = 0 \quad (3-45)$$

The solutions are e^{ikx} and e^{-ikx} , or linear combinations of these, such as $\cos kx$ and $\sin kx$. The trouble with all of them is that they are not square-integrable, since

$$\int_{-\infty}^{\infty} dx |Ae^{ikx} + Be^{-ikx}|^2$$

diverges for all values of A and B . There are three ways of getting around this difficulty.

- (a) One way is to consider the particle as confined in a very large box. The difficulty with a box extending from $-a/2$ to $+a/2$ is that the wave functions must

vanish at the boundaries, and it is therefore impossible to construct wave functions of the type $\exp(\pm ikx)$. A way of getting around this is to replace the boundary condition $\psi(\pm a/2) = 0$ by *periodic boundary conditions* expressed in the form

$$\psi(x) = \psi(x + a) \quad (3-46)$$

For a free particle described by e^{ikx} this implies that k cannot be continuous, but must satisfy

$$e^{ika} = 1 \quad (3-47)$$

which means that $k = 2n\pi/a$ or that $p = 2n\pi\hbar/a$. If the box is large, then the n -values must be correspondingly large. The normalization is one that corresponds to one “period” of width a , so that

$$N^2 \int_0^a dx |e^{ikx}|^2 = N^2 a = 1 \quad (3-48)$$

Thus the solutions corresponding to an energy $\frac{1}{2m} (2n\pi\hbar/a)^2$ will be linear combinations of $e^{ikx}/\sqrt{a}$ and $e^{-ikx}/\sqrt{a}$. Since a is arbitrary, we expect the answer to all sensible questions¹ to be independent of a . This is indeed the case.

(b) We may work with wave packets. A solution of the form

$$\psi(x) = e^{ikx} \quad (3-49)$$

is a special case of (3-43) with

$$\phi(p) = \sqrt{2\pi\hbar} \delta(p - \hbar k) \quad (3-50)$$

that is, an infinitely peaked momentum-space distribution. Suppose we replace this limiting $\phi(p)$ by a very sharply peaked function $\sqrt{2\pi\hbar} g(p - \hbar k)$. Then e^{ikx} will be replaced by

$$\begin{aligned} \psi(x) &= \int_{-\infty}^{\infty} dp e^{ipx/\hbar} g(p - \hbar k) \\ &= e^{ikx} \int_{-\infty}^{\infty} dq e^{iqx/\hbar} g(q) \end{aligned} \quad (3-51)$$

which is a plane wave, e^{ikx} , multiplied by a very broad function of x . We may make this function so broad that it is essentially constant over the region of physical interest. The uncertainty in the momentum will now be of the order of magnitude $\hbar/(\text{size of } x\text{-packet})$, and if the denominator is of macroscopic size, this uncertainty is negligible. We thus satisfy the mathematical requirements without changing any of the physics. The wave-packet description is actually the one that is closest to what really happens physically, since any way of preparing the initial state, for example, firing an electron gun, can never, in practice, create an exact momentum eigenstate.

¹An example of a question that is not sensible is “How long does it take a particle to travel from $x = 0$ to $x = a/2$?”

- (c) We can approach the problem by recognizing that the difficulty in normalization stems from the fact that for a wave function like e^{ikx} , the particle is not confined to any region of space, so that the probability of finding it anywhere is zero. If we do not ask questions that involve the probability of finding the particle in any finite region, no problems arise. One way of avoiding the normalization difficulty is to deal with the probability current, or *flux*

$$j(x) = \frac{\hbar}{2im} \left[\psi^*(x) \frac{d\psi(x)}{dx} - \frac{d\psi^*(x)}{dx} \psi(x) \right] \quad (3-52)$$

discussed in Chapter 2. For a wave function $Ce^{ipx/\hbar}$, the flux is $|C|^2 p/m$; for the wave function $Ce^{-ipx/\hbar}$, the flux is $-|C|^2 p/m$. If we note that for a one-dimensional problem, the flux of particles with a density of 1 particle/m, moving with velocity $v = p/m$ is just v —that is, the number crossing a point $x = x_0$ per second—we see that $|C|^2$ represents the density of particles per m. Thus (3-41) represents particles with a density of $1/2\pi\hbar$ per m.

In three dimensions, with

$$u_p(\mathbf{r}) = Ce^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \quad (3-53)$$

the flux will be $|C|^2 \mathbf{p}/m$, and this corresponds to a flow of particles, with density $|C|^2$ per m^3 crossing a unit area perpendicular to $\mathbf{p}$, when the particles are moving with velocity $\mathbf{v} = \mathbf{p}/m$ (Fig. 3-4).

Degeneracy

The energy eigenvalue equation (3-45) has two independent solutions, e^{ikx} and e^{-ikx} , equivalently, the pair of real solutions $\cos kx$ and $\sin kx$ is also independent. Whichever pair we choose, we notice that in contrast to the problem of a particle in a box, there are *two* solutions that have the same energy associated with them. This is an example of something that happens quite frequently: *there may be more than one independent eigenfunction that corresponds to the same eigenvalue of a hermitian operator*. When this occurs, we have a *degeneracy*.

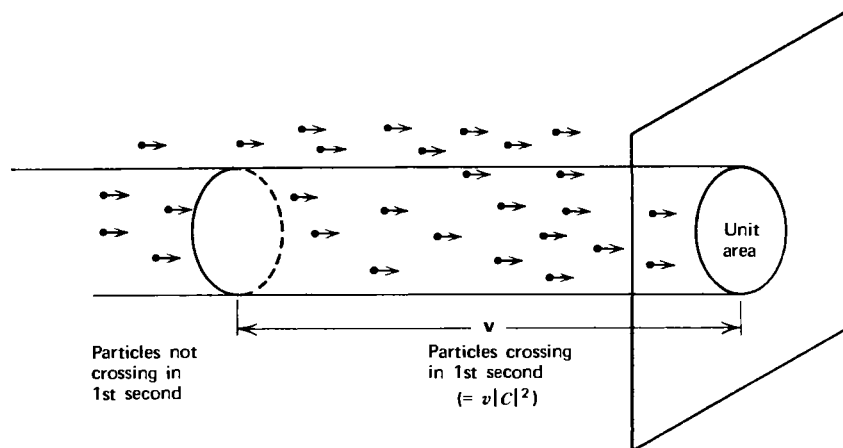


Figure 3-4 The relation between velocity of particles and flux—that is, number of particles crossing a unit area perpendicular to velocity, per unit time.

In the two cases that we have earlier, the two solutions are orthogonal:

$$\begin{aligned} \int_{-\infty}^{\infty} dx (e^{-ikx})^* e^{ikx} &= \int_{-\infty}^{\infty} dx e^{2ikx} = 0 \\ \int_{-\infty}^{\infty} dx \sin kx \cos kx &= 0 \end{aligned} \quad (3-54)$$

for $k \neq 0$. It is always possible to make linear combinations such that this is true. Independent of any degeneracy, eigenfunctions corresponding to a given energy $E = \hbar^2 k^2 / 2m$ are orthogonal to eigenfunctions corresponding to a *different energy*.

What distinguishes the two degenerate eigenfunctions? For the set (e^{-ikx}, e^{ikx}) , the difference is that they are eigenfunctions of the momentum operator

$$p_{\text{op}} e^{\pm ikx} = \frac{\hbar}{i} \frac{d}{dx} e^{\pm ikx} = \pm \hbar k e^{\pm ikx} \quad (3-55)$$

The two degenerate energy eigenfunctions are distinguished by the fact that they are simultaneously eigenfunctions of the momentum operator, with eigenvalues $\pm p$.

We could, instead have chosen $\cos kx$ and $\sin kx$ as degenerate solutions. We expect that there must be something more basic than the fact that they are different functions that distinguishes them. The important difference turns out to be that one of them is *even* under the interchange $x \rightarrow -x$ and the other one is *odd* under that interchange. This leads us to yet another operator of interest, called *the parity operator*.

3-6 PARITY

The eigenfunctions for the free particle ($\sin kx, \cos kx$) as well as the eigenfunctions for a particle in a box extending from $-a/2$ to $+a/2$, are either even or odd under the interchange $x \rightarrow -x$. Suppose we consider the particle in a box that is symmetric about the $x = 0$ point, and suppose that the initial state $\psi(x)$ is even in x . This means that $\psi(x)$ must be of the form

$$\psi(x) = \sum_n A_n \sqrt{\frac{2}{a}} \cos \frac{n\pi x}{a} \quad (3-56)$$

where the sum extends over $n = 1, 3, 5, \dots$. At a later time this wave function becomes

$$\psi(x, t) = \sum_n A_n \sqrt{\frac{2}{a}} \cos \frac{n\pi x}{a} e^{-iE_n t / \hbar} \quad (3-57)$$

Thus the wave function is still even in x at a later time. The same holds for a wave function that is initially odd. Thus for our box, symmetrically centered about $x = 0$, *evenness* and *oddness* are properties that are time independent. We may say that evenness and oddness are *constants of the motion*. Since any constant of the motion is of interest to us, we formalize the discussion somewhat.

We do this by introducing the *parity operator* P , whose rule of operation is to reflect $x \rightarrow -x$. Thus for any wave function $\psi(x)$, we have

$$P\psi(x) = \psi(-x) \quad (3-58)$$

Let

$$\psi(x, 0) = \psi(-x, 0) \equiv \psi^{(+)}(x) \quad (3-64)$$

The time development is given by

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = H\psi(x, t) \quad (3-65)$$

If we operate with P on this equation, we get

$$i\hbar \frac{\partial}{\partial t} P\psi(x, t) = PH\psi(x, t) \quad (3-66)$$

Under the special circumstances that

$$PH\psi(x, t) = HP\psi(x, t) \quad (3-67)$$

which holds when H is even under $x \rightarrow -x$ —that is, when $V(x)$ is an even function (since d^2/dx^2 is even)—we have

$$i\hbar \frac{\partial}{\partial t} [P\psi(x, t)] = H[P\psi(x, t)] \quad (3-68)$$

Hence

$$\psi^{(+)}(x, t) = \frac{1}{2}(1 + P)\psi(x, t) \quad (3-69)$$

and

$$\psi^{(-)}(x, t) = \frac{1}{2}(1 - P)\psi(x, t) \quad (3-70)$$

separately obey the Schrödinger equation, and do not mix, if the initial state is even (or odd). The condition for the time independence of parity only holds if

$$(PH - HP)\psi(x, t) = 0 \quad (3-71)$$

for all possible states—that is, if the operators P and H commute

$$[P, H] = 0 \quad (3-72)$$

This important condition will be seen to be quite general: *any operator that does not have an explicit time dependence and that commutes with the Hamiltonian H is a constant of the motion*. In particular, if the potential changes with time—that is, we have $V(x, t)$ —then the energy itself is not a constant of the motion, just as in classical mechanics. Note that when V depends on t , the separation of the equation into an equation for the time dependence and an energy eigenvalue equation is not possible.

In summary, what differentiates the degenerate eigenfunctions is that they are simultaneous eigenfunctions of another hermitian operator. Both the operators p_{op} and P have the property that they commute with the Hamiltonian $p_{\text{op}}^2/2m$ in this problem. We shall show later that this is a necessary condition for the existence of simultaneous eigenfunctions. For example, p_{op} and P do not commute [since $(\hbar/i)(d/dx)$ changes sign under $x \rightarrow -x$], and therefore the eigenfunctions of one of the operators cannot all be simultaneous eigenfunctions of the other.

We have learned an enormous amount about quantum mechanics from the two simple problems that we have considered. We shall return to these matters in later chapters and generalize them. In Chapter 4 we will again consider some very simple problems. However, there we will concentrate not on the mathematical features, but rather on the physical systems that they are simple models of.

PROBLEMS

1. You are given the following operators:

(a) $O_1\psi(x) = x^3\psi(x)$; (b) $O_2\psi(x) = x(d/dx)\psi(x)$; (c) $O_3\psi(x) = \lambda\psi^*(x)$;
 (d) $O_4\psi(x) = e^{\psi(x)}$; (e) $O_5\psi(x) = [d\psi(x)/dx] + a$; (f) $O_6\psi(x) = \int_{-\infty}^x dx' (\psi(x')x')$

Which of these are linear operators?

2. Solve the eigenvalue problem

$$O_6\psi(x) = \lambda\psi(x)$$

What values of the eigenvalue λ lead to square-integrable eigenfunctions? (*Hint*: Differentiate both sides of the equation with respect to x .)

3. Calculate the following commutators: (a) $[O_2, O_6]$; (b) $[O_1, O_2]$. The procedure is to calculate $[A, B]$ by expressing $A(B\psi) - B(A\psi)$ in the form $C\psi$.

4. Calculate

$$\Delta x = \sqrt{\langle x^2 \rangle}$$

for the $u_n(x)$ given in (3-21). Using $\langle p^2 \rangle$ given by (3-26), calculate

$$\Delta p \Delta x$$

It is characteristic that for the higher states the uncertainty increases with n .

5. Consider an electron of mass $m = 0.9 \times 10^{-30}$ kg in an infinite box of dimension $a = 10^{-9}$ m.
 (a) What is the energy difference between the ground state and the first excited state? Express your answer in eV.
 (b) Suppose the transition from the state $n = 2$ to the state $n = 1$ is accompanied by the emission of a photon, as given by the Bohr rule. What is the wavelength of the emitted photon?
6. Consider an electron in a macroscopic box of size $a = 2$ cm.
 (a) What value of n corresponds to an energy of 1.5 eV?
 (b) What is the difference in energy between the state n and $n + 1$ in that energy region?
7. Consider an infinite box of unknown width. In transitions between neighboring values of n , photons of various energies are emitted. It is found that the largest wavelength of the various photons seen is 450×10^{-9} m. Use this information to determine a , the width of the infinite box.
8. Consider an infinite box extending from $x = -a/2$ to $x = a/2$. Can you, without actually solving the Schrödinger equation but by looking at the shape of the wave functions, immediately give a general formula for the energy eigenvalues of an infinite box that extends for $x = -a/2$ to $x = 0$? (*Hint*: Where do the eigenfunctions have *nodes*?)
9. A particle is known to be localized in the left half of a box with sides at $x = \pm a/2$, with wave function

$$\psi(x) = \begin{cases} \sqrt{\frac{2}{a}} & -\frac{a}{2} < x < 0 \\ 0 & 0 < x < \frac{a}{2} \end{cases}$$

- (a) Will the particle remain localized at later times?
 (b) Calculate the probability that an energy measurement yields the ground state energy; the energy of the first excited state.

10. Consider the wave function of Problem 9.
- (a) Calculate the probability that an energy measurement yields the energy eigenvalue associated with a particular value of n .
 - (b) Use the fact that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{8}$$

to show that the probabilities add up to one.

11. The eigenfunctions for a potential of the form

$$V(x) = \begin{cases} \infty & x < 0; \quad x > a \\ 0 & 0 < x < a \end{cases}$$

are of the form

$$u_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}$$

Suppose a particle in the preceding potential has an initial normalized wave function of the form

$$\psi(x, 0) = A \left(\frac{\sin \pi x}{a} \right)^5$$

- (a) What is the form of $\psi(x, t)$?
 - (b) Calculate A without doing the integral $\int d\theta \sin^{10} \theta$.
 - (c) What is the probability that an energy measurement yields E_3 , where $E_n = n^2 \pi^2 \hbar^2 / 2ma^2$?
{Hint: Expand $([e^{i\theta} - e^{-i\theta}]/2i)^5$ in a power series $e^{5i\theta} + \dots - e^{-5i\theta}$ and recombine into a series of terms involving $\sin 5\theta$ and so on.}
12. A particle is in the ground state of a box with sides at $x = \pm a$. Very suddenly the sides of the box are moved to $x = \pm b$ ($b > a$). What is the probability that the particle will be found in the ground state for the new potential? What is the probability that it will be found in the first excited state? In the latter case, the simple answer has a simple explanation. What is it?
13. Repeat the calculation given in example (3-6) for a particle initially in the n th eigenstate. Show that the corresponding probability is given by

$$\frac{2n^2\pi}{\hbar a^3} \frac{1 - (-1)^n \cos p a / \hbar}{[(p/\hbar)^2 - (n\pi/a)^2]^2}$$

Sketch the distribution. Show that it conforms with the uncertainty relation, and that the result is in agreement with the correspondence principle when n is large.

14. A particle in free space is initially in a wave packet described by

$$\psi(x) = \left(\frac{\alpha}{\pi} \right)^{1/4} e^{-\alpha x^2/2}$$

- (a) What is the probability that its momentum is in the range $(p, p + dp)$?
 - (b) What is the expectation value of the energy? Can you give a rough argument based on the “size” of the wave function and the uncertainty principle, for why the answer should be roughly what it is?
15. The wave function for a particle is given by

$$\psi(x) = A e^{ikx} + B e^{-ikx}$$

What flux does this represent?

16. What is the flux associated with a particle described by the wave function

$$\psi(x) = u(x) e^{ikx}$$

where $u(x)$ is a real function?

17. Consider the eigenfunctions for a box with sides at $x = \pm a$. Without working out the integral, prove that the expectation value of the quantity

$$x^2 p^3 + 3xp^3x + p^3x^2$$

vanishes for all the eigenfunctions.

18. Prove that the parity operator, defined by

$$P\psi(x) = \psi(-x)$$

is a hermitian operator. Also prove that the eigenfunctions of P , corresponding to the eigenvalues $+1$ and -1 , are orthogonal.

19. Use parity arguments to show that in Example 3-5 the A_n for n even must all vanish.

Chapter 4

One-Dimensional Potentials

In this chapter we solve some simple problems of one-dimensional motion. They are of interest because they illustrate some nonclassical effects, and because many physical situations are effectively one-dimensional even though we live in a three-dimensional world.

4-1 THE POTENTIAL STEP

For this problem we take (Fig. 4-1) the form of $V(x)$ to be

$$\begin{aligned} V(x) &= 0 & x < 0 \\ &= V_0 & x > 0 \end{aligned} \quad (4-1)$$

The Schrödinger equation

$$-\frac{\hbar^2}{2m} \frac{d^2 u(x)}{dx^2} + V(x) u(x) = E u(x) \quad (4-2)$$

takes the form

$$\frac{d^2 u(x)}{dx^2} + \frac{2m}{\hbar^2} [E - V(x)] u(x) = 0 \quad (4-3)$$

We write, as usual,

$$\frac{2mE}{\hbar^2} = k^2 \quad (4-4)$$

and we also introduce

$$\frac{2m(E - V_0)}{\hbar^2} = q^2 \quad (4-5)$$

The most general solution of (4-3) for $x < 0$, where $V(x) = 0$, is

$$u(x) = e^{ikx} + R e^{-ikx} \quad (4-6)$$

This corresponds to a net flux moving in the positive x -direction, of magnitude

$$\begin{aligned} j &= \frac{\hbar}{2im} \left(u^* \frac{du}{dx} - \frac{du^*}{dx} u \right) = \frac{\hbar}{2im} [(e^{-ikx} + R^* e^{ikx})(ike^{ikx} - ikR e^{-ikx}) - \text{c.c.}] \\ &= \frac{\hbar k}{m} (1 - |R|^2) \end{aligned} \quad (4-7)$$

We may view e^{ikx} with flux $\hbar k/m$ as an *incoming wave*. If there were no potential, we could choose e^{ikx} as the solution for all x , so that we attribute R to the presence of the potential. This potential gives rise to a reflected wave, $R e^{-ikx}$, with a reflected flux $\hbar k|R|^2/m$.

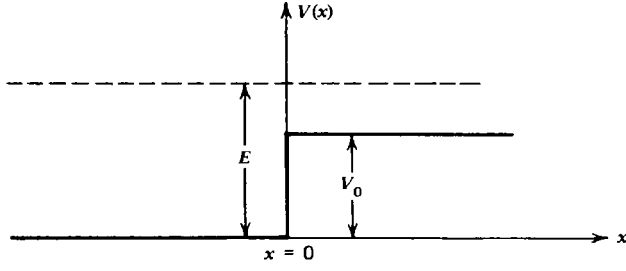


Figure 4-1 The potential step.

For $x > 0$, we write the solution

$$u(x) = T e^{iqx} \quad (4-8)$$

The most general solution for $x > 0$ is a linear combination of e^{iqx} and e^{-iqx} , but a term involving the latter would describe a wave coming from $+\infty$ in the negative direction, and with the “experiment” in which only a wave from the left is sent in, the only wave on the right can be a transmitted wave. The flux corresponding to (4-8) is

$$j = \frac{\hbar q}{m} |T|^2 \quad (4-9)$$

Since there is no time dependence in the problem, the conservation law (3-12) implies that $j(x)$ is independent of x . Hence the flux on the left must be equal to the flux on the right; that is, we expect that

$$\frac{\hbar k}{m} (1 - |R|^2) = \frac{\hbar q}{m} |T|^2 \quad (4-10)$$

The continuity of the wave function at $x = 0$ implies that

$$1 + R = T \quad (4-11)$$

obtained by matching the two solutions at $x = 0$. In spite of the fact that the potential is discontinuous, the slope of the wave function is also continuous, as can be seen by integrating (4-3) from $-\varepsilon$ to $+\varepsilon$ (with ε arbitrarily small and positive) and using the continuity of the wave function:

$$\begin{aligned} \left(\frac{du}{dx} \right)_{\varepsilon} - \left(\frac{du}{dx} \right)_{-\varepsilon} &= \int_{-\varepsilon}^{\varepsilon} dx \frac{d}{dx} \frac{du}{dx} \\ &= \int_{-\varepsilon}^{\varepsilon} dx \frac{2m}{\hbar^2} [V(x) - E] u(x) = 0 \end{aligned} \quad (4-12)$$

We note, for future reference, that if the potential contains a term like $\lambda \delta(x - a)$, then integration of the equation from $a - \varepsilon$ to $a + \varepsilon$ gives

$$\begin{aligned} \left(\frac{du}{dx} \right)_{a+\varepsilon} - \left(\frac{du}{dx} \right)_{a-\varepsilon} &= \frac{2m}{\hbar^2} \int_{a-\varepsilon}^{a+\varepsilon} dx \lambda \delta(x - a) u(x) \\ &= \frac{2m}{\hbar^2} \lambda u(a) \end{aligned} \quad (4-13)$$

The continuity of the derivative for our potential implies that

$$ik(1 - R) = iqT \quad (4-14)$$

We can therefore solve (4-11) and (4-14) for R and T to obtain

$$\begin{aligned} R &= \frac{k - q}{k + q} \\ T &= \frac{2k}{k + q} \end{aligned} \quad (4-15)$$

From this we can calculate the reflected and transmitted fluxes:

$$\begin{aligned} \frac{\hbar k}{m} |R|^2 &= \frac{\hbar k}{m} \left(\frac{k - q}{k + q} \right)^2 \\ \frac{\hbar q}{m} |T|^2 &= \frac{\hbar k}{m} \frac{4kq}{(k + q)^2} \end{aligned} \quad (4-16)$$

We note the following:

1. In contrast to classical mechanics, according to which a particle going over a potential step would slow down (to conserve energy) but would never be reflected, here we do have a certain fraction of the incident particles reflected. This is, of course, a consequence of the wave properties of the particle; partial reflection of light from an interface between two media is a familiar phenomenon.
2. With the help of (4-16) we can easily check that the conservation law (4-10) is indeed satisfied.
3. For $E \gg V_0$ —that is, for $q \rightarrow k$ from below—the ratio of the reflected flux to the incident flux—that is, $|R|^2$ —approaches zero. This agrees with intuition, which tells us that at very high energies, the presence of the step is but a small perturbation on the propagation of the wave.
4. If the energy E is less than V_0 , then q becomes imaginary. If we note that, now the solution for $x > 0$ must be of the form

$$u(x) = T e^{-|q|x} \quad (4-17)$$

so as not to blow up at $+\infty$, we see that in this case,

$$|R|^2 = \left(\frac{k - i|q|}{k + i|q|} \right) \left(\frac{k - i|q|}{k + i|q|} \right)^* = 1 \quad (4-18)$$

Thus, as in classical mechanics, there is now total reflection. Note, however, that

$$T = \frac{2k}{k + i|q|} \quad (4-19)$$

does not vanish, and a part of the wave penetrates into the forbidden region. This penetration phenomenon again is characteristic of waves, and we shall see a little later that it permits a “tunneling” through barriers that would totally block particles in a classical description. There is no flux to the right, since $j(x)$ vanishes for a real solution even if the coefficient in front of it is taken to be complex. The phenomenon of total reflection is mathematically identical to what happens to light when it strikes an interface between two media of different refractive indices (going from the larger to the smaller value of n) at an angle

larger than the critical angle. The light undergoes total internal reflection, but there is an exponentially decaying electromagnetic field that penetrates into the forbidden region.

5. It is a peculiarity of the idealized sharp potential that the expressions for R and T can be written in terms of E and $(E - V_0)$ only, independent of $\hbar$, which could apparently be set equal to zero. This appears to present us with a paradox, since we associate $\hbar = 0$ with the classical limit. Does this mean that for such a potential one could get partial reflection in the classical limit? There is no paradox, since a condition for the classical limit to hold is that the de Broglie wavelength ($\lambda = 2\pi\hbar/p$) be small compared with the relevant dimensions of the system. In this example, the dimension that is relevant is the size of the region in which the potential changes from zero to V_0 , which is zero in the limiting case. Thus for this idealized potential, there never is a classical region. If we were to round off the potential, then at sufficiently high energies the condition for classical behavior would indeed obtain, but as already noted, at sufficiently high energies there is no reflection.

4-2 THE POTENTIAL WELL

We next consider the potential (Fig. 4-2)

$$\begin{aligned} V(x) &= 0 & x < -a \\ &= -V_0 & -a < x < a \\ &= 0 & a < x \end{aligned} \quad (4-20)$$

We again write

$$k^2 = \frac{2mE}{\hbar^2} \quad (4-21)$$

and

$$q^2 = \frac{2m(E + V_0)}{\hbar^2} \quad (4-22)$$

We can immediately write down the solutions

$$\begin{aligned} u(x) &= e^{ikx} + R e^{-ikx} & x < -a \\ u(x) &= A e^{iqx} + B e^{-iqx} & -a < x < a \\ u(x) &= T e^{ikx} & a < x \end{aligned} \quad (4-23)$$

These correspond to an incoming flux $\hbar k/m$ from the left, a reflected flux $\hbar k|R|^2/m$, and a transmitted flux $\hbar k|T|^2/m$ to the right. Inside the well there are waves going in both direc-

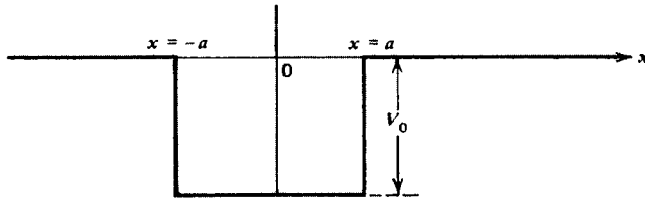


Figure 4-2 The potential well.

tions because of the reflections at both discontinuities at $\pm a$. According to flux conservation we should get

$$\frac{\hbar k}{m} (1 - |R|^2) = \frac{\hbar q}{m} (|A|^2 - |B|^2) = \frac{\hbar k}{m} |T|^2 \quad (4-24)$$

We again write down the conditions for the continuity of $u(x)$ and $du(x)/dx$ at $x = a$ and $x = -a$. It is useful to note that these conditions can be combined into a single set of conditions requiring the continuity of

$$\frac{1}{u(x)} \frac{du(x)}{dx}$$

at $x = a$ and $x = -a$. In this form we see that the overall normalization plays no role in determining the shape (though not the scale) of the wave functions. Multiplying $u(x)$ by a constant will not change the ratio. In the present situation we get

At $x = -a$

$$\frac{iqAe^{-iqa} - iqBe^{iqa}}{Ae^{-iqa} + Be^{iqa}} = \frac{ike^{-ika} - ikRe^{ika}}{e^{-ika} + Re^{ika}} \quad (4-25a)$$

At $x = a$

$$\frac{iqAe^{iqa} - iqBe^{-iqa}}{Ae^{iqa} + Be^{-iqa}} = \frac{ikTe^{ika}}{Te^{ika}} = ik \quad (4-25b)$$

A little algebra is needed to eliminate A and B . What remains are expressions for R and T . These are

$$R = i e^{-2ika} \frac{(q^2 - k^2) \sin 2qa}{2kq \cos 2qa - i(q^2 + k^2) \sin 2qa} \quad (4-26)$$

$$T = e^{-2ika} \frac{2kq}{2kq \cos 2qa - i(q^2 + k^2) \sin 2qa}$$

Again, if $E \gg V_0$, there is practically no reflection, since $q^2 - k^2 \ll 2kq$, and as $E \rightarrow 0$, the transmission goes to zero. There is an item of special interest: in the special case that $\sin 2qa = 0$ —that is, for the positive energies given by

$$E = -V_0 + \frac{n^2 \pi^2 \hbar^2}{8ma^2} \quad n = 1, 2, 3, \dots \quad (4-27)$$

there is no reflection. This is actually a model of what happens in the scattering of low energy electrons (0.1 eV) by noble gas atoms—for example, neon and argon—in which there is anomalously large transmission. The effect, first observed by Ramsauer and Townsend, is described as a *transmission resonance*. A more accurate discussion must, of course, involve three-dimensional considerations. In wave language, the effect is due to a destructive interference between the wave reflected at $x = -a$ and the wave reflected once, twice, thrice, . . . , at the edge $x = a$. The resonance condition $2qa = n\pi$, which may be written in the form

$$\lambda = \frac{2\pi}{q} = \frac{4a}{n} \quad (4-28)$$

is just the one that describes the Fabry-Perot interferometer in optics.

4-3 THE POTENTIAL BARRIER

We now consider

$$\begin{aligned} V(x) &= 0 & x < -a \\ &= V_0 & -a < x < a \\ &= 0 & a < x \end{aligned} \quad (4-29)$$

We will limit our discussion to energies such that $E < V_0$ —that is, energies such that no penetration of the barrier would occur in classical physics (Fig. 4-3). Inside the barrier we have the equation

$$\frac{d^2 u(x)}{dx^2} + \frac{2m}{\hbar^2} (E - V_0) u(x) = 0$$

that is, with

$$\begin{aligned} \kappa^2 &= -\frac{2m(E - V_0)}{\hbar^2} \\ \frac{d^2 u(x)}{dx^2} - \kappa^2 u(x) &= 0 \end{aligned} \quad (4-30)$$

The general solution

$$u(x) = Ae^{-\kappa x} + Be^{\kappa x} \quad |x| < a \quad (4-31)$$

is to be matched onto

$$\begin{aligned} u(x) &= e^{ikx} + Re^{-ikx} & x < -a \\ &= Te^{ikx} & x > a \end{aligned} \quad (4-32)$$

Actually we need not go through the trouble of solving this since the results can be read off from (4-26) with the substitution

$$q \rightarrow i\kappa = i\sqrt{(2m/\hbar^2)(V_0 - E)} \quad (4-33)$$

Thus, for example,

$$T = e^{-2ika} \frac{2k\kappa}{2k\kappa \cosh 2\kappa a - i(k^2 - \kappa^2) \sinh 2\kappa a} \quad (4-34)$$

and this implies that

$$|T|^2 = \frac{(2k\kappa)^2}{(k^2 + \kappa^2)^2 \sinh^2 2\kappa a + (2k\kappa)^2} \quad (4-35)$$

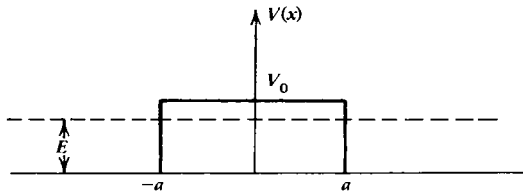


Figure 4-3 Potential barrier. Energy is such that a classical particle would be totally reflected by the barrier.

When $\kappa a \gg 1$, $\sinh 2\kappa a \rightarrow \frac{1}{2} e^{2\kappa a}$, and the expression for $|T|^2$ simplifies to

$$|T|^2 \rightarrow \left(\frac{4\kappa k}{\kappa^2 + k^2} \right)^2 e^{-4\kappa a} \quad (4-36)$$

There is transmission, even though the energy lies below the top of the barrier. This is a wave phenomenon, and in quantum mechanics it is also one exhibited by particles. This *tunneling* of a particle through a barrier is frequently encountered, and we shall discuss some applications later.

We pause to analyze an apparent difficulty. The wave function does not vanish inside the barrier, and thus there appears to be some probability of finding a particle with negative kinetic energy. How can this make sense? We look to the uncertainty relation to remove an apparent paradox that arises from *too classical* a description of the process. An experiment to study the particle inside the potential barrier must be able to localize it with an accuracy

$$\Delta x \ll 2a \quad (4-37)$$

This measurement will transfer to the particle momentum, with an uncertainty

$$\Delta p \gg \hbar/2a \quad (4-38)$$

which corresponds to a transfer of energy

$$\Delta E \gg \hbar^2/8ma^2 \quad (4-39)$$

In order to observe the negative kinetic energy, this uncertainty must be much less than $|E - V_0|$, so that

$$\frac{\hbar^2 \kappa^2}{2m} \gg \Delta E \gg \frac{\hbar^2}{8ma^2} \quad (4-40)$$

which implies that $2\kappa a \gg 1$. Under these circumstances, the quantity to be measured, $|T|^2$ is, as seen in (4-36), vanishingly small. For example for $\kappa a = 10$, $e^{-4\kappa a} \approx 10^{-18}$.

The approximate expression for the ratio of transmitted flux to incident flux, $|T|^2$, is an extremely sensitive function of the width of the barrier, and of the amount by which the barrier exceeds the incident energy, since

$$\kappa a = \left[\frac{2ma^2}{\hbar^2} (V_0 - E) \right]^{1/2} \quad (4-41)$$

In general, the barriers that occur in physical phenomena are not square, and to discuss some applications, we must first obtain an approximate expression for the transmission coefficient $|T|^2$ through an irregularly shaped barrier. The proper way to do this, given the fact that there is no exact solution available for most potentials, is through the Wentzel–Kramers–Brillouin (WKB) approximation technique.¹ Our discussion will be less mathematical.

We observe that (4-36) consists of a product of a *pre-factor* and a much more rapidly varying exponential factor. Thus we may isolate this behavior by writing, again approximately

$$\ln |T|^2 = (\text{const}) - (2\kappa)(2a)$$

¹The WKB approximation is briefly discussed in Supplement 4-A. [www.wiley.com/college/gasiorowicz]

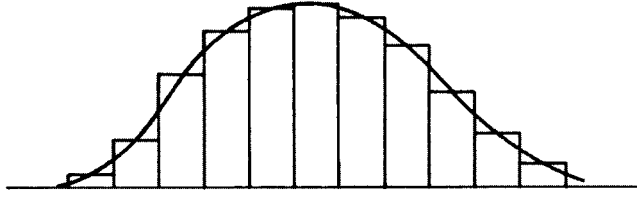


Figure 4-4 Approximation of a smooth barrier by a juxtaposition of square potential barriers.

in which the energy dependent factor 2κ multiplies the *width* of the barrier $2a$. The procedure we adopt is to treat a smooth curved barrier as a juxtaposition of square barriers, as shown in Fig. 4-4. If the transmission probability for each barrier is small, the overall probability is the product of the individual ones. Equivalently the overall transmission coefficient is a product of the transmission coefficients of the individual barriers. In effect, when most of the flux is reflected by a single barrier, the transmission through each “slice” is an independent, improbable event. We may therefore write, approximately

$$\begin{aligned} \ln |T|^2 &= \sum_{\text{slices}} \ln |T_{\text{slice}}|^2 = -2 \sum_n \Delta x_n \langle \kappa \rangle_n \\ &= -2 \int_{\text{barrier}} dx \sqrt{2m(V(x) - E)/\hbar^2} \end{aligned} \quad (4-42)$$

In the “slices” Δx_n is the width and $\langle \kappa \rangle$ is the average value of κ for that slice of the barrier. In the last step, a limit of infinitely narrow barriers was taken. It is clear from the expression that the approximation is least accurate near the *turning points*, where the energy and potential are nearly equal, since there (4-36) is not a good approximation to (4-35). It is also important that $V(x)$ be a slowly varying function of x , since otherwise the approximation of a smooth barrier by a stack of square ones is only possible if the latter are narrow. However, there (4-36) is again a poor approximation. A proper treatment using the WKB approximation includes a discussion of the behavior near the turning points. For our purposes it is a fair approximation to write

$$|T|^2 = Ce^{-2 \int dx \sqrt{2m(V(x) - E)/\hbar^2}} \quad (4-43)$$

where the integration is over the region in which the square root is real.

4-4 AN EXAMPLE OF TUNNELING

We illustrate tunneling by considering *cold emission*. We learned from our discussion of the photoelectric effect that electrons need a minimum energy W (the *work function*) to escape from a metal. We may therefore view the most easily liberated electrons as trapped in a large box, with a potential barrier of height W . The photoelectric effect is not the only way in which electrons can be removed from the metal. They can also be removed by heating the metal, and they can be removed at room temperature by the application of an external electric field $\mathcal{E}$. This phenomenon is described as *cold emission*. It occurs because the external field changes the potential seen by the electron from a macroscopically wide barrier of height W to one described by $W - e\mathcal{E}x$, where x is the distance from the wall of the box (Fig. 4-5b). The change creates a barrier of finite width, and electrons can tunnel through it. We define the most easily removed electrons as having energy $E = 0$. This then gives the transmission coefficient as

$$|T|^2 = Ce^{-2 \int_0^a dx \sqrt{2m(W - e\mathcal{E}x)/\hbar^2}} \quad (4-44)$$

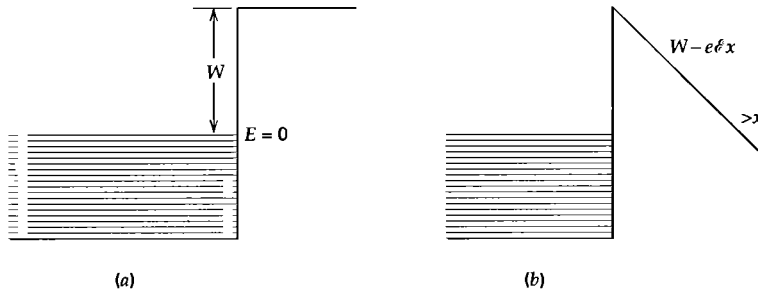


Figure 4-5 (a) Electronic levels in metal. We set the energy of the highest level $E = 0$. (b) Schematic sketch of how the potential is altered by an external electric field.

where a is the width of the barrier given by

$$a = \frac{W}{e\mathcal{E}} \quad (4-45)$$

The integral in (4-44) is easily evaluated. It leads to the result

$$|T|^2 \cong Ce^{-\frac{4\sqrt{2}}{3}\sqrt{mWa^2\hbar^2}} \quad (4-46)$$

The *Fowler-Nordheim formula*, as (4-46) is called, describes emission only qualitatively. One effect, which is easily included, is the additional attraction of the electron to its image charge, which acts to pull it back to the plate. The other effect, much harder to handle, is that there are surface imperfections, which change the electric field locally, and since $\mathcal{E}$ appears in the exponent (in a) this can make a large difference.

In the last 25 years a version of cold emission has found an important application in the *scanning tunneling microscope* (STM). This tool is based on the extreme sensitivity to the distance in cold emission. The current of electrons induced by an electric potential difference between a metal surface and a very sharp tungsten tip above the surface varies exponentially with the distance to the surface, and this allows for very detailed studies of the topography of the surface. The functioning of the STM depends on some new concepts and also on several technological advances.

First, in order to get a 1-Å (or less) resolution, radiation of the order of 10 keV is required by the usual diffraction limit. This does not allow for nondestructive studies of a surface. An electron microscope requires electrons of the order of 150 eV, which is still too large for the scanning of surfaces. As was pointed out by J. A. O'Keefe in 1956, it is possible to bypass the diffraction limit with a new type of microscope. In such a microscope light would shine through a tiny hole onto the object directly in front of the hole, and the transmitted or reflected light would be affected by the object. In such a microscope it is the size of the hole that determines its resolution. The STM operates on this principle. The tungsten tip, to which the potential is applied, acts as the "hole."

A potential difference of 10 V acting across a spacing of about 5 Å gives rise to very large electric fields, and these fields "pull out" a few atoms from the tip to sharpen it. For example, a tip machined to have a radius of 1000 Å would quite naturally be able to resolve lateral distances to 45 Å because of the sensitivity of the current to the distance from the sampled surface. The tiny atomic protuberances improve this resolution to 1–2 Å.

The usefulness of the STM depends on the ability of the experimenter to keep the distance from the surface, or alternatively, the current constant. This can now be achieved by

means of ceramic (piezoelectric) supports that expand or contract when electric fields are applied to them. Their location can be determined to very high precision by means of interferometers, and thus a constant distance between the tip and the scanned surface can be maintained.

The STM has been used for a study of surfaces of metals and some semiconductors. More recently atoms and molecules have been moved around on surfaces, either by sliding them by means of the tip or by lifting them off the surface by the tip and depositing them at a different location.

Another example of tunneling is given in Supplement 4-B [www.wiley.com/college/gasiorowicz], where an application to nuclear decay with the emission of α -particles is discussed.

4-5 BOUND STATES IN A POTENTIAL WELL

In addition to the solutions for $E > 0$ discussed in the section on the potential well, there are also solutions for $E < 0$. This is only possible if the potential well is *attractive*; that is, $V_0 > 0$ in (4-20). In classical physics too, negative energies appear, as for the Keplerian orbits with a potential $-GmM/r$. What is different about quantum mechanics is that the negative energy states have a discrete, rather than a continuous spectrum. Let us write

$$\frac{2mE}{\hbar^2} = -\alpha^2 \quad (4-47)$$

The solutions outside the well that are bounded at infinity are:

For $x < -a$

$$u(x) = C_1 e^{\alpha x}$$

For $a < x$

$$u(x) = C_2 e^{-\alpha x} \quad (4-48)$$

Since we are dealing with real functions, it is more convenient to write the solution inside the well in the form

$$u(x) = A \cos qx + B \sin qx \quad -a < x < a \quad (4-49)$$

Note that

$$q^2 = \frac{2m}{\hbar^2} (V_0 - |E|) > 0 \quad (4-50)$$

The potential under consideration is symmetric under reflection in the origin, when $x \rightarrow -x$. The kinetic energy is manifestly unaltered by this change, which means that the Hamiltonian is unchanged under $x \rightarrow -x$. We may therefore classify solutions by their *parity* in addition to their energy eigenvalues. Thus we may separately look at solutions that are *even* in x , and solutions that are *odd* in x . For the even solutions $C_2 = C_1$ while for the odd solutions, $C_2 = -C_1$. As Fig. 4-6 shows, it is sufficient to match

$$\frac{1}{u(x)} \frac{du(x)}{dx}$$

at $x = a$ only, since the same conditions will obtain when the matching is done at $x = -a$. This is clearly a consequence of the symmetry under reflection.

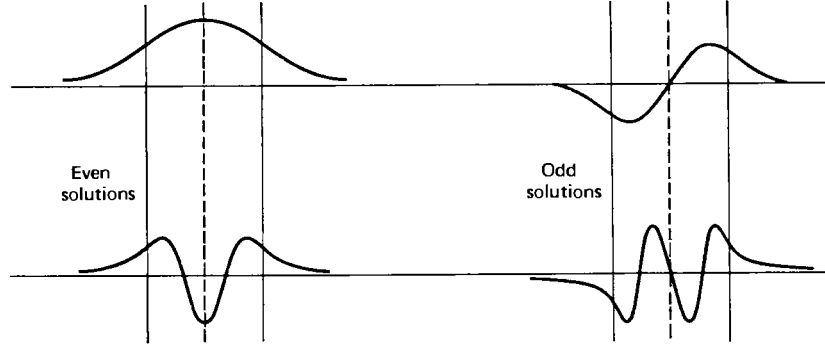


Figure 4-6 Solutions for a discrete spectrum in an attractive potential well.

For the even solutions, $B = 0$ and we have

$$-\alpha = -q \frac{\sin qa}{\cos qa}$$

that is,

$$\alpha = q \tan qa \quad (4-51)$$

For the odd solutions, $A = 0$ and we get

$$-\alpha = q \frac{\cos qa}{\sin qa}$$

that is,

$$\alpha = -q \cot qa \quad (4-52)$$

Let us examine these separately.

(a) The even solutions:

With the notation

$$\begin{aligned} \lambda &= \frac{2mV_0a^2}{\hbar^2} \\ y &= qa \end{aligned} \quad (4-53)$$

the relation (4-51) reads

$$\frac{\sqrt{\lambda - y^2}}{y} = \tan y \quad (4-54)$$

If we plot $\tan y$ and $\sqrt{\lambda - y^2}/y$ as functions of y (Fig. 4-7), the points of intersection determine the eigenvalues. These form a discrete set. The larger λ is, the further the curves for $\sqrt{\lambda - y^2}/y$ go; that is, *when the potential is deeper and/or broader, there are more bound states*. The figure also shows that no matter how small λ is, there will always be at least one bound state. This is characteristic of one-dimensional attractive potentials, and is *not* true for three-dimensional potentials, which behave much more like the odd-solution problem that we will discuss later. As λ becomes large, the eigenvalues tend to become equally spaced in y , with the intersection points given approximately by

$$y \simeq (n + \frac{1}{2})\pi \quad n = 0, 1, 2, \dots \quad (4-55)$$

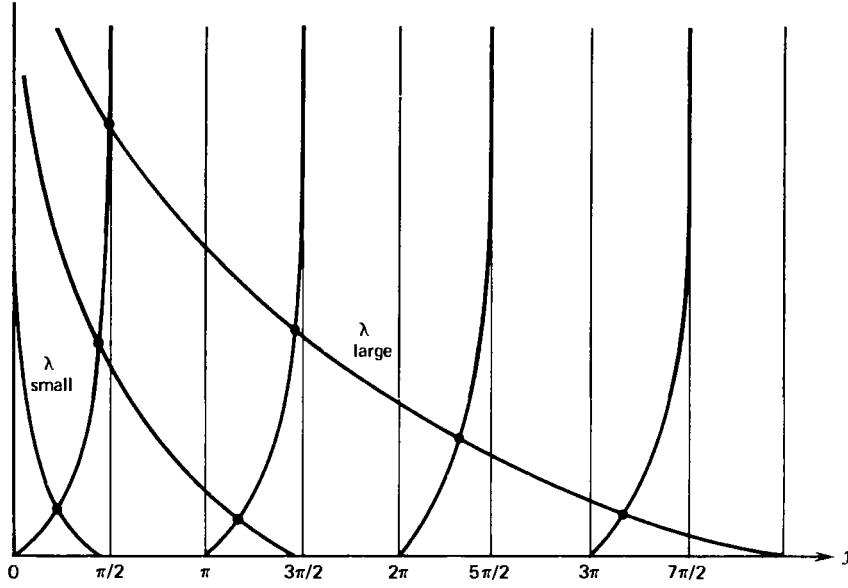


Figure 4-7 Location of discrete eigenvalues for even solutions in a square well. The rising curves represent $\tan y$; the falling curves are $\sqrt{\lambda - y^2}/y$ for different values of λ .

This is just the eigenvalue condition for the even solutions of the infinite box centered at the origin [eigenfunctions given in (3-28)]. This is as might be expected, since for the deep-lying states in the potential, the fact that it is not really infinitely deep does not matter very much.

(b) The odd solutions:

Here the eigenvalue condition reads

$$\frac{\sqrt{\lambda - y^2}}{y} = -\cot y \quad (4-56)$$

Since $-\cot y = \tan(\pi/2 + y)$, the plot in Fig. 4-8 is the same as in Fig. 4-7 with the tangent curves shifted by $\pi/2$. The large λ behavior is more or less the same, with (4-55) replaced by

$$y \approx n\pi \quad n = 1, 2, 3, \dots \quad (4-57)$$

In contrast to the even solutions, there will only be an intersection if $\sqrt{\lambda - \pi^2/4} > 0$, that is, if

$$\frac{2mV_0a^2}{\hbar^2} \geq \frac{\pi^2}{4} \quad (4-58)$$

The odd solutions all vanish at $x = 0$, and hence the bound-state problem for the odd solutions will be the same as for the potential well shown in Fig. 4-9 since in the latter, the condition $u(0) = 0$ would be imposed. We shall see that such conditions are imposed on wave functions in the three-dimensional world.

The detailed calculations that we have carried out support the qualitative understanding of the reason for discrete eigenvalues. These arise because we require the wave functions to vanish at infinity. We see this graphically in Fig. 4-10. The even-parity ground-state wave function inside the potential, of the form $\cos qx$, must tie on continuously to a falling

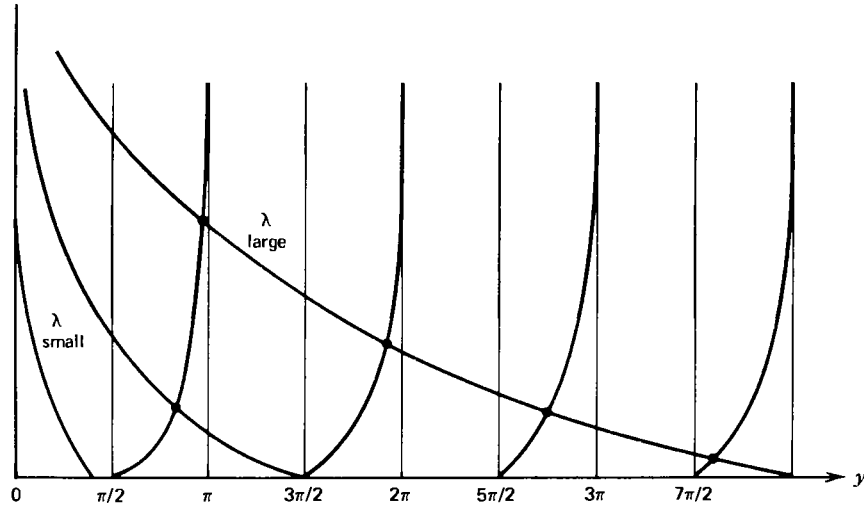


Figure 4-8 Location of discrete eigenvalues for odd solutions in a square well. The rising curves represent $-\cot y$; the falling curves are $\sqrt{\lambda - y^2}/y$ for different values of λ . Note that there is no eigenvalue for $\lambda < (\pi/2)^2$.

exponential $e^{-\alpha|x|}$ with $\alpha^2 = 2mE_B/\hbar^2$. Large binding means a rapidly falling exponential. Since $q^2 = (2mV_0/\hbar^2 - \alpha^2)$, a large binding energy means that q^2 is small; that is, the wave function is quite flat: thus matching is impossible. As we reduce the trial E_B , the exponential falls less steeply and the wave function inside curves more, so that at some point the matching (continuous slope) becomes possible. If the value of α is lowered beyond this point, the outside curve is too flat to match the more curved inside wave function. For the first excited state, with odd parity, the wave function vanishes at the origin, so that it can only tie onto a falling exponential if it has a chance to turn over inside the potential. The condition that it turns over just enough to tie to a straight line ($\alpha = 0$) is that $\sin qa = 1$, so that $qa = \pi/2$, which corresponds to the condition expressed in (4-58).

The sketches we used to describe the wave functions for the bound states can be used to give some insight for more complicated potentials. Let us take, for example, a single well shown in Fig. 4-11(a) and then a double well, consisting of two wells identical to those in Fig. 4-11(a), separated by a distance somewhat larger than the width of each well (Fig. 4-11b). Consider first the lowest, *even* energy state in the single well. We see that to match slopes at the edges of the well the wave function on the inside must match the falling, exponential curves on both sides outside the well. This determines the rate at which the exponential falls, and which thus determines the binding energy. If we look at

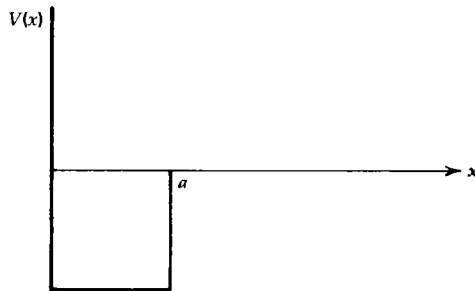


Figure 4-9 Equivalent potential for odd solutions of a square well bound state problem.

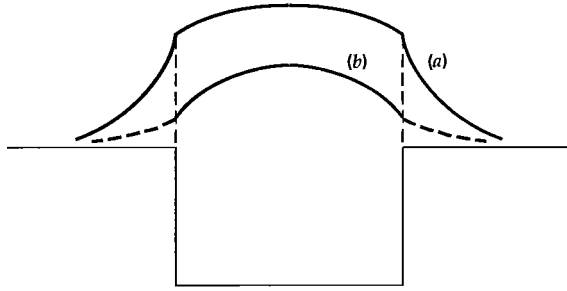


Figure 4-10 (a) Mismatch with E_B too large. (b) Mismatch with E_B too small.

the double well, the wave function must be symmetric about the midpoint between the two wells. However, this does not prevent the wave function inside the well on the right, for example, from dropping quite a bit more on the far right side, and still connecting to a fairly flat curve between the wells, as shown in Fig. 4-11(b). Thus a double well can bind a particle much more strongly than a single well, not because two wells are somewhat stronger than one well, but rather because one can maintain an almost constant wave function between the two potentials. This means that the bound particle has a high probability of being anywhere between the two wells.

For the first excited state (*odd* under reflection), the wave function must go through zero at the midpoint, so that it matches on to a flatter falling curve at the outer edges of the wells. This implies a smaller binding, as expected. When the wells are very far apart, then, roughly speaking, a particle bound on the right side does not know that there is a well on the left, and the same is true for the particle when it is bound to the well on the left. If the wave functions for the two cases are $u_R(x)$ and $u_L(x)$, then the symmetric wave function is

$$u_{\text{even}}(x) = \frac{1}{N_e} (u_R(x) + u_L(x)) \quad (4-59)$$

while the antisymmetric wave function, corresponding to the first excited state, is

$$u_{\text{odd}}(x) = \frac{1}{N_o} (u_R(x) - u_L(x)) \quad (4-60)$$

The normalization factors differ slightly because of the nonvanishing overlap of the $u_R(x)$ and $u_L(x)$. If the wells were so far apart that there was no overlap, then the normalization constants would each be $1/\sqrt{2}$. Because of the small overlap, the eigenvalues are not

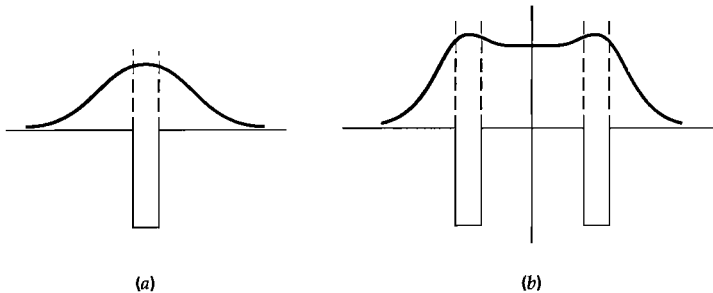


Figure 4-11 (a) Wave function for the lowest bound state in a single square well potential. (b) Wave function for lowest bound state in a double well. Note that the wave functions at the inner edges of the wells do not have to fall off as rapidly to satisfy the continuity condition for $u(x)$ and $du(x)/dx$, allowing for a large wave function between the wells and a steeper drop-off at the outer edges.

quite the same. Thus the $u_R(x)$ and the $u_L(x)$ do not represent stationary states. For example, the wave function of a particle localized on the right will be approximately given by

$$u(x) \approx C(u_{\text{even}}(x) + u_{\text{odd}}(x)) \quad (4-61)$$

If we start with a particle “on the right” at time $t = 0$, then at a later time the wave function will be

$$\begin{aligned} u(x, t) &= C(u_{\text{even}}(x)e^{-iE_e t/\hbar} + u_{\text{odd}}(x)e^{-iE_o t/\hbar}) \\ &= Ce^{-iE_e t/\hbar}(u_{\text{even}}(x) + u_{\text{odd}}(x)e^{-i(E_o - E_e)t/\hbar}) \\ &\equiv Ce^{-iE_e t/\hbar}(u_{\text{even}}(x) + u_{\text{odd}}(x)e^{-i\Delta E t/\hbar}) \end{aligned} \quad (4-62)$$

Now after a time t such that $e^{-i\Delta E t/\hbar} = -1$ —that is, after a time

$$t = \frac{\hbar\pi}{\Delta E} = \frac{\hbar\pi}{E_o - E_e} \quad (4-63)$$

the wave function becomes one that is approximately localized on the left side. The particle thus oscillates between the two wells with a period $2t$ —that is, with an angular frequency

$$\omega = \frac{\Delta E}{\hbar} \quad (4-64)$$

This will be illustrated in detail when we discuss the delta-function potentials in the next section. For now we point out that the above considerations apply to a variety of physical systems. A good example is the ammonia molecule. NH_3 has the tetrahedral shape shown in Figure 4-12 with the three H nuclei forming a triangular base, and the N nucleus at one of the two symmetrically located apexes of the pyramid. This location is determined by the condition that the energy is a minimum, which implies that in the region between the two minima there is a potential barrier. The two-well problem models this, and we might expect two nearly degenerate states, with the N nucleus oscillating between the two minima. This is actually observed, with a frequency of 1.500×10^{11} radians/s. This frequency has been determined with very high accuracy by placing the ammonia samples in resonant cavities and tuning those to match the oscillation frequency of the ammonia molecule. The way this is done is to prepare a beam of ammonia molecules heated in an oven (at 300 K). The beam contains molecules in the ground as well as in the nearby vibrational excited state. When the beam is passed through an electrostatic separator, the different electric dipole moments of the two states allow the separation of a beam of pure “upper state” molecules into a microwave cavity resonator. When the electric field in the cavity oscillates with a frequency that just matches the transition frequency ω , then transitions to the ground state are induced. The energy emitted in the form of photons can be detected.

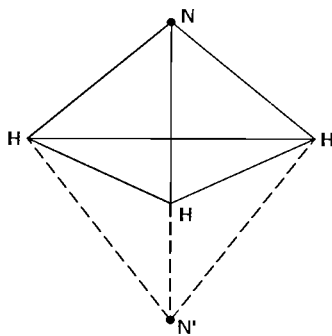


Figure 4-12 Schematic representation of localized sites of nuclei in an ammonia molecule. The nitrogen nucleus oscillates between N and N'.

4-6 DELTA FUNCTION POTENTIALS

We consider a potential $V(x)$ whose spatial behavior is given by $\delta(x)$. Since $\delta(x)$ has the dimensions of a reciprocal length, it is convenient to write for the attractive potential.

$$V(x) = -\frac{\hbar^2 \lambda}{2ma} \delta(x) \quad (4-65)$$

We have introduced a , an arbitrary quantity with dimensions of a length, so that λ is a dimensionless quantity that we use to characterize the strength of the potential. The equation to be solved is, when $E < 0$,

$$\frac{d^2 u(x)}{dx^2} - \kappa^2 u(x) = -\frac{\lambda}{a} \delta(x) u(x) \quad (4-66)$$

where $\kappa^2 = 2m|E|/\hbar^2$. The solution everywhere, except at $x = 0$, must satisfy the equation $d^2 u/dx^2 - \kappa^2 u = 0$, and if it is to vanish at $x \rightarrow \pm\infty$, we must have

$$u(x) = \begin{cases} e^{-\kappa x} & x > 0 \\ e^{\kappa x} & x < 0 \end{cases} \quad (4-67)$$

The coefficients in front are the same (and here chosen to be unity—we can normalize afterwards) because of the continuity of the wave function. The derivative of the wave function is no longer continuous. As argued before (Eq. 4-13), we have

$$\left(\frac{du}{dx}\right)_{x=0+} - \left(\frac{du}{dx}\right)_{x=0-} = -\frac{\lambda}{a} u(0) \quad (4-68)$$

The last relation gives the eigenvalue condition

$$-\kappa - \kappa = -\frac{\lambda}{a}$$

that is,

$$\kappa = \frac{\lambda}{2a} \quad (4-69)$$

The double delta function potential is more interesting because it gives a quick way to study the properties of a narrow deep double well, such as shown in Figure 4-13. Let

$$(2m/\hbar^2) V(x) = -\frac{\lambda}{a} [\delta(x-a) + \delta(x+a)] \quad (4-70)$$

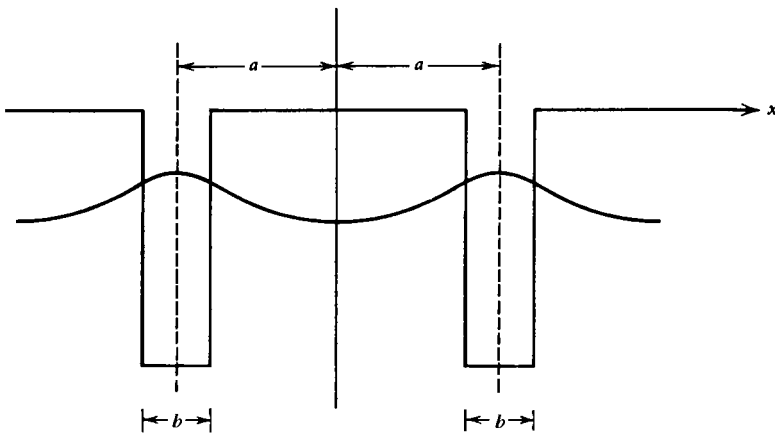


Figure 4-13 Double one-dimensional potential well. The shape of the wave function for a bound state is sketched in. We consider the limiting case in (4-70).

(Here a in the description of the magnitude of the potential is no longer an arbitrary length, but is connected to the form of the potential.) Because the potential is symmetric under the interchange $x \rightarrow -x$, we expect that there will be solutions of definite parity, and we will first consider the even solutions.

1. For the even solution we write

$$\begin{aligned} u(x) &= e^{-\kappa x} & x > a \\ &= A \cosh \kappa x & a > x > -a \\ &= e^{\kappa x} & x < -a \end{aligned} \quad (4-71)$$

and continuity of the wave function gives

$$e^{-\kappa a} = A \cosh \kappa a \quad (4-72)$$

Because of the symmetry, it is sufficient to apply the discontinuity condition for the derivative at $x = a$. Nothing new will come of the application at $x = -a$. We get

$$-\kappa e^{-\kappa a} - \kappa A \sinh \kappa a = -\frac{\lambda}{a} e^{-\kappa a} \quad (4-73)$$

and the eigenvalue condition is, upon eliminating A ,

$$\tanh \kappa a = \frac{\lambda}{\kappa a} - 1 \quad (4-74)$$

The preceding may be rewritten in the form

$$\frac{\lambda}{\kappa a} - 1 = \frac{e^{\kappa a} - e^{-\kappa a}}{e^{\kappa a} + e^{-\kappa a}} = \frac{1 - e^{-2\kappa a}}{1 + e^{-2\kappa a}}$$

from which we get

$$e^{-2\kappa a} = \frac{2\kappa a}{\lambda} - 1 \quad (4-75)$$

For λ large we must have $2\kappa a$ slightly larger than λ . If we try $2\kappa a = \lambda + \varepsilon$, then, to order ε (4.75) reads $e^{-\lambda} \approx \varepsilon/\lambda$ so that

$$2\kappa a = \lambda + \lambda e^{-\lambda} \quad (4-76)$$

We can see that for the even solution there is always a single bound state. Figure 4-14 shows a plot of the eigenvalue equation (4-74) with $y = \kappa a$. As the figure shows, there is only one intersection point of the curve $\tanh y$ with $(\lambda/y) - 1$. It is obvious that when $y = \lambda$, the right side is zero, whereas $\tanh y > 0$. Thus the inter-

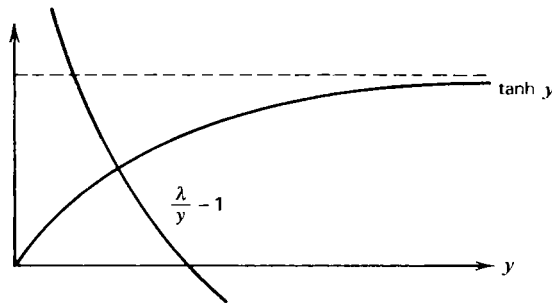


Figure 4-14 Solution of the eigenvalue condition $\tanh y = \lambda/y - 1$.

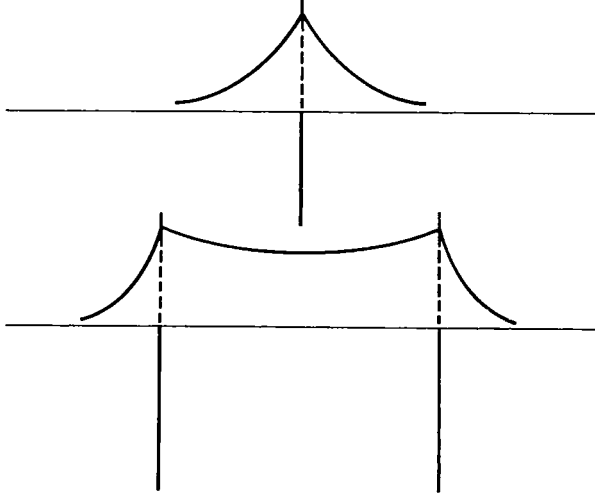


Figure 4-15 Bound state wave functions for single and double delta function attractive potentials. The strength of the potential is measured by the discontinuity in the slopes of the wave function at the potentials. It is the same in all three cases, and shows that for the double potential, a much steeper fall-off to the left and to the right is possible.

section point occurs for $y < \lambda$. On the other hand, since $\tanh y < 1$, we must have $(\lambda/y) < 2$ at the intersection point; that is,

$$\kappa > \frac{\lambda}{2a} \quad (4-77)$$

If we compare this with (4-69), we see that the energy for the double well is a *larger negative number*; that is, the energy for the double potential is lower. Note that this is not because somehow the strength of a pair of potentials is larger than that of a single potential, as might be the case if one compared an electron bound to two protons with an electron bound to one proton. The larger binding is there because, as Figure 4-15 indicates, it is easier to accommodate a sharply dropping exponential to a symmetric function (here $\cosh x$) with a discontinuity in slope as given, than it is to accommodate it to an equally sharply dropping exponential on the other side of the potential. In the real world, a single electron bound to two protons separated by a small distance will have a lower energy than a single proton plus a hydrogen atom far away, even though in the first case there is a more effective repulsion between the protons. Again it is the way in which the wave function can accommodate itself to the geometrical situation that is the dominant effect.

2. The odd solution will have the form

$$\begin{aligned} u(x) &= e^{-\kappa x} & x > a \\ &= A \sinh \kappa x & a > x > -a \\ &= -e^{\kappa x} & x < -a \end{aligned} \quad (4-78)$$

Again, because of the antisymmetry, it is sufficient to apply the conditions at $x = a$, say. Continuity of the wave function gives

$$A \sinh \kappa a = e^{-\kappa a} \quad (4-79)$$

and the discontinuity equation reads

$$-\kappa e^{-\kappa a} - \kappa A \cosh \kappa a = -\frac{\lambda}{a} e^{-\kappa a} \quad (4-80)$$

Combining the two yields the eigenvalue condition

$$\coth \kappa a = \frac{\lambda}{\kappa a} - 1 \quad (4-81)$$

This equation can be rewritten in the form

$$\frac{\lambda}{\kappa a} - 1 = \frac{e^{\kappa a} + e^{-\kappa a}}{e^{\kappa a} - e^{-\kappa a}} = \frac{1 + e^{-2\kappa a}}{1 - e^{-2\kappa a}}$$

Just as for the ground-state eigenvalue, we can find the solution for $\lambda \gg 1$. It is given by changing the sign of $e^{-2\kappa a}$ so that

$$2\kappa a = \lambda - \lambda e^{-\lambda} \quad (4-82)$$

We can show that the odd solution has at most one bound state. Figure 4-16 shows a plot of the reciprocal of the eigenvalue equation—that is, of

$$\tanh \kappa a = \left(\frac{\lambda}{\kappa a} - 1 \right)^{-1}$$

with κa denoted by y .

There will only be an intersection if the slope of the former at the origin is larger than that of the second—that is, if

$$\lambda > 1 \quad (4-83)$$

At $y = \lambda/2$ the term $(\lambda/y - 1)^{-1}$ is already at 1, so that the intersection had to occur for $y < \lambda/2$; that is,

$$\kappa < \frac{\lambda}{2a} \quad (4-84)$$

Thus the odd solution, if there is a bound state, is less strongly bound than the even solution. The wave function, which has to go through zero, is forced to be steep between the wells, and thus can only accommodate to a less rapidly falling exponential. Depending on the size of λ , there may or may not exist an odd parity bound state.

For large λ we can use our solutions for the energies to obtain

$$\begin{aligned} \omega &= \frac{1}{\hbar} \frac{\hbar^2}{2ma^2} (-(\kappa_o a)^2 - (-\kappa_e a)^2) \\ &= \frac{\hbar}{8ma^2} ((\lambda + \lambda e^{-\lambda})^2 - (\lambda - \lambda e^{-\lambda})^2) \\ &= \frac{\hbar \lambda^2}{2ma^2} e^{-\lambda} \end{aligned} \quad (4-85)$$

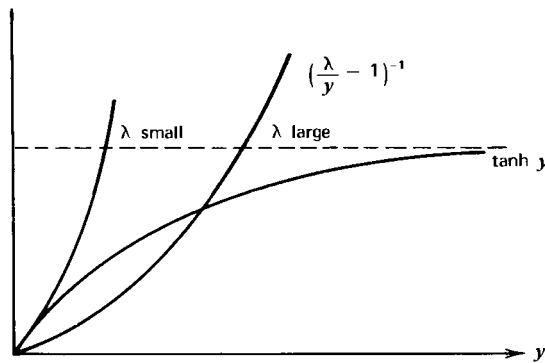


Figure 4-16 Solution of the eigenvalue condition $\tanh y = (\lambda/y - 1)^{-1}$.

4-7 THE HARMONIC OSCILLATOR

As our last example we consider the harmonic oscillator (Fig. 4-17). In contrast to the examples dealt with until now, the differential equation that needs to be solved is not so trivial, and one reason for discussing this problem is to learn something about the techniques for solving such equations.

The classical Hamiltonian is of the form

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 \quad (4-86)$$

so that the eigenvalue equation is

$$-\frac{\hbar^2}{2m} \frac{d^2 u(x)}{dx^2} + \frac{1}{2}kx^2 u(x) = Eu(x) \quad (4-87)$$

We introduce the frequency of the oscillator

$$\omega = \sqrt{k/m} \quad (4-88)$$

write

$$\varepsilon = \frac{2E}{\hbar\omega} \quad (4-89)$$

and change variables to

$$y = \sqrt{\frac{m\omega}{\hbar}} x \quad (4-90)$$

to finally get the simpler form of the equation

$$\frac{d^2 u}{dy^2} + (\varepsilon - y^2) u = 0 \quad (4-91)$$

All quantities that appear are dimensionless.

For any eigenvalue ε , as $y^2 \rightarrow \infty$, the term involving ε is negligible, and we must therefore require that $u(y)$ asymptotically satisfy the equation

$$\frac{d^2 u_0(y)}{dy^2} - y^2 u_0(y) = 0 \quad (4-92)$$

We multiply by $2du_0/dy$, which allows us to rewrite this in the form

$$\frac{d}{dy} \left(\frac{du_0}{dy} \right)^2 - y^2 \frac{d}{dy} (u_0^2) = 0$$

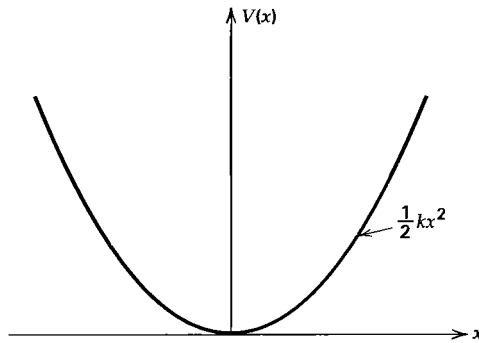


Figure 4-17 Harmonic oscillator potential.

or, equivalently,

$$\frac{d}{dy} \left[\left(\frac{du_0}{dy} \right)^2 - y^2 u_0^2 \right] = -2y u_0^2 \quad (4-93)$$

This simplifies a great deal if we neglect the term on the right side of the equation. We assume that this can be done, and then check that the assumption was correct. If we drop the right side, we find that

$$\frac{du_0}{dy} = (C + y^2 u_0^2)^{1/2} \quad (4-94)$$

where C is a constant of integration. Since both $u_0(y)$ and du_0/dy must vanish at infinity, we must have $C = 0$. Thus

$$\frac{du_0}{dy} = \pm y u_0 \quad (4-95)$$

whose solution, acceptable at infinity, is

$$u_0(y) = e^{-y^2/2} \quad (4-96)$$

We can now check that $2y u_0^2 = 2y e^{-y^2}$ is indeed negligible compared with

$$\frac{d}{dy} (y^2 u_0^2) = \frac{d}{dy} (y^2 e^{-y^2}) \simeq -2y^3 e^{-y^2} \quad (4-97)$$

for large y . If we now introduce a new function $h(y)$, such that

$$u(y) = h(y) e^{-y^2/2} \quad (4-98)$$

then the differential equation is easily seen to take the form

$$\frac{d^2 h(y)}{dy^2} - 2y \frac{dh(y)}{dy} + (\varepsilon - 1) h(y) = 0 \quad (4-99)$$

This may not seem like much of a simplification, but we have accounted for the behavior at infinity, and we can now look at the behavior near $y = 0$. Let us attempt a power series expansion

$$h(y) = \sum_{m=0}^{\infty} a_m y^m \quad (4-100)$$

When this is inserted into the equation, we find that the coefficients of y^m satisfy the recursion relation

$$(m+1)(m+2) a_{m+2} = (2m - \varepsilon + 1) a_m \quad (4-101)$$

Thus, given a_0 and a_1 , the even and odd series can be generated separately. That they do not mix is a consequence of the invariance of the Hamiltonian under reflections. For arbitrary ε , we find that for large m (say, $m > N$)

$$a_{m+2} \simeq \frac{2}{m} a_m \quad (4-102)$$

This means that the solution is approximately

$$h(y) = (\text{a polynomial in } y) + a_N \left[y^N + \frac{2}{N} y^{N+2} + \frac{2^2}{N(N+2)} y^{N+4} + \frac{2^3}{N(N+2)(N+4)} y^{N+6} + \dots \right]$$

where, for simplicity, we have only taken the even solution. The series may be written in the form

$$a_N y^2 \left(\frac{N}{2} - 1 \right)! \left[\frac{(y^2)^{N/2-1}}{(N/2-1)!} + \frac{(y^2)^{N/2}}{(N/2)!} + \frac{(y^2)^{N/2+1}}{(N/2+1)!} + \dots \right]$$

If we choose $N = 2k$ for convenience, the series takes the form

$$\begin{aligned} y^2(k-1)! \left[\frac{(y^2)^{k-1}}{(k-1)!} + \frac{(y^2)^k}{k!} + \frac{(y^2)^{k+1}}{(k+1)!} + \dots \right] \\ = y^2(k-1)! \left[e^{y^2} - \left\{ 1 + y^2 + \frac{(y^2)^2}{2!} + \dots + \frac{(y^2)^{k-2}}{(k-2)!} \right\} \right] \end{aligned}$$

which is of the form of a polynomial + a constant $\times y^2 e^{y^2}$. When this is inserted into (4-98), we get a solution that does not vanish at infinity. An acceptable solution can be found if the recursion relation (4-101) terminates—that is, if

$$\varepsilon = 2n + 1 \quad (4-103)$$

For that particular value of ε the recursion relations yield

$$a_{2k} = (-2)^k \frac{n(n-2) \cdots (n-2k+4)(n-2k+2)}{(2k)!} a_0 \quad (4-104)$$

and

$$a_{2k+1} = (-2)^k \frac{(n-1)(n-3) \cdots (n-2k+3)(n-2k+1)}{(2k+1)!} a_1 \quad (4-105)$$

Thus the results are:

1. There are discrete, equally spaced eigenvalues. Equation (4-103) translates into

$$E = \hbar\omega(n + \frac{1}{2}); n = 0, 1, 2, \dots \quad (4-106)$$

a form that looks familiar, since the relation between energy and frequency is the same as that discovered by Planck for the radiation field modes. This is no accident, since a decomposition of the electromagnetic field into normal modes is essentially a decomposition into harmonic oscillators that are decoupled.

2. The polynomials $h(y)$ are, except for normalization constants, the Hermite polynomials $H_n(y)$, whose properties can be found in any number of textbooks on mathematical physics. We limit ourselves to the following outline of their properties:

$H_n(y)$ satisfy the differential equation

$$\frac{d^2 H_n(y)}{dy^2} - 2y \frac{dH_n(y)}{dy} + 2nH_n(y) = 0 \quad (4-107)$$

They satisfy the following recursion relations

$$H_{n+1} - 2yH_n + 2nH_{n-1} = 0 \quad (4-108)$$

$$H_{n+1} + \frac{dH_n}{dy} - 2yH_n = 0 \quad (4-109)$$

Also,

$$\sum_n H_n(y) \frac{z^n}{n!} = e^{2zy - z^2} \quad (4-110)$$

and

$$H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2} \quad (9-111)$$

The normalization of the Hermite polynomials is such that

$$\int_{-\infty}^{\infty} dy e^{-y^2} H_n(y)^2 = 2^n n! \sqrt{\pi} \quad (4-112)$$

We list a few of the Hermite polynomials here:

$$\begin{aligned} H_0(y) &= 1 \\ H_1(y) &= 2y \\ H_2(y) &= 4y^2 - 2 \\ H_3(y) &= 8y^3 - 12y \\ H_4(y) &= 16y^4 - 48y^2 + 12 \\ H_5(y) &= 32y^5 - 160y^3 + 120y \end{aligned}$$

The orthogonality of eigenfunctions corresponding to different values of n is easily established. The eigenvalue equations

$$\frac{d^2 u_n}{dx^2} = \frac{mk}{\hbar^2} x^2 u_n - \frac{2mE_n}{\hbar^2} u_n$$

and

$$\frac{d^2 u_l^*}{dx^2} = \frac{mk}{\hbar^2} x^2 u_l^* - \frac{2mE_l}{\hbar^2} u_l^*$$

when multiplied by u_l^* and u_n , respectively, and the second equation is subtracted from the first one, yields

$$\frac{d}{dx} \left(u_l^* \frac{du_n}{dx} - \frac{du_l^*}{dx} u_n \right) = \frac{2m}{\hbar^2} (E_l - E_n) u_l^* u_n$$

When this equation is integrated over x from $-\infty$ to $+\infty$, the left-hand side vanishes, since the eigenfunctions and their derivatives vanish at $x = \pm\infty$. Thus

$$(E_l - E_n) \int_{-\infty}^{\infty} dx u_l^*(x) u_n(x) = 0 \quad (4-114)$$

which means that the eigenfunctions for which $E_l \neq E_n$ are orthogonal. The reason for the importance of the harmonic oscillator in quantum mechanics, as in classical mechanics, is that any small perturbation of a system from its equilibrium state will give rise to small oscillations, which are ultimately decomposable into normal modes—that is, independent oscillators.

3. As (4-106) shows, even the lowest state has some energy, the *zero-point energy*. Its presence is a purely quantum mechanical effect, and can be interpreted in terms of the uncertainty principle. It is the zero-point energy that is responsible for the fact that helium does not “freeze” at extremely low temperatures, but remains liquid down to temperatures of the order of 10^{-3} K, at normal pressures. The fre-

quency ω is larger for lighter atoms, which is why the effect is not seen for nitrogen, say. It also depends on detailed features of the interatomic forces, which is why liquid hydrogen does freeze.

Figure 4-18 shows the shapes of the lowest six eigenfunctions.

Another class of one-dimensional potentials of physical interest are periodic potentials, which satisfy the condition

$$V(x) = V(x + a)$$

Such potentials lead to *band structure* in the energy spectrum—that is, continuous values of allowed energies separated by gaps. This is a rather space- and time-consuming project, and we leave it to Supplement 4-C [www.wiley.com/college/gasiorowicz].

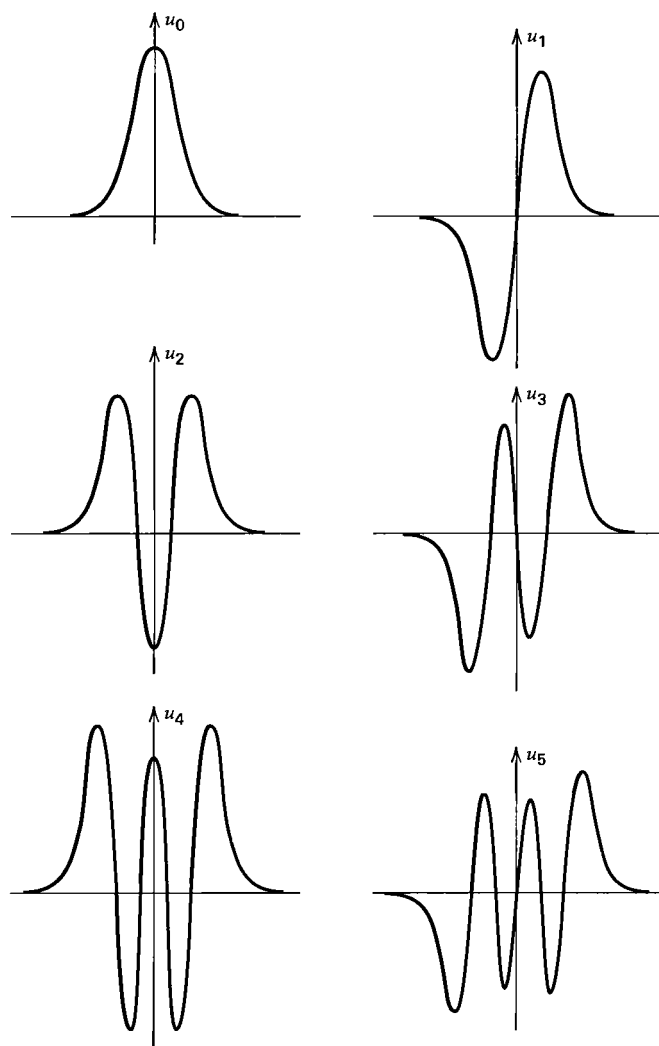
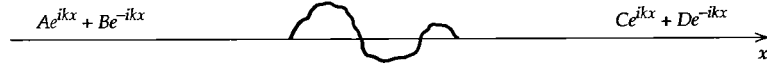


Figure 4-18 The shapes of the first six eigenfunctions.

PROBLEMS

1. Consider an arbitrary potential localized on a finite part of the x -axis. The solutions of the Schrödinger equation to the left and to the right of the potential region are given by



$$Ae^{ikx} + Be^{-ikx} \qquad \qquad \qquad Ce^{ikx} + De^{-ikx}$$

respectively. Show that if we write

$$\begin{aligned} C &= S_{11}A + S_{12}D \\ B &= S_{21}A + S_{22}D \end{aligned}$$

that is, relate the “outgoing” waves to the “ingoing” waves by

$$\begin{pmatrix} C \\ B \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A \\ D \end{pmatrix}$$

then the following relations hold

$$\begin{aligned} |S_{11}|^2 + |S_{21}|^2 &= 1 \\ |S_{12}|^2 + |S_{22}|^2 &= 1 \\ S_{11}S_{12}^* + S_{21}S_{22}^* &= 0 \end{aligned}$$

Use this to show that the *matrix*

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

and its transpose are unitary. (*Hint:* Use flux conservation and the possibility that A and D are arbitrary complex numbers.)

2. Calculate the elements of the scattering matrix, S_{11} , S_{12} , S_{21} , and S_{22} for the potential

$$\begin{aligned} V(x) &= 0 & x < -a \\ &= V_0 & -a < x < a \\ &= 0 & x > a \end{aligned}$$

and show that the general conditions proved in Problem 1 are indeed satisfied.

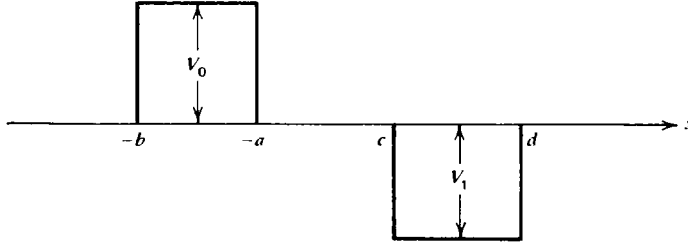
3. The elements $S_{11} \cdots S_{22}$ are functions of k . Show that

$$\begin{aligned} S_{11}(-k) &= S_{11}^*(k) \\ S_{22}(-k) &= S_{22}^*(k) \\ S_{12}(-k) &= S_{21}^*(k) \end{aligned}$$

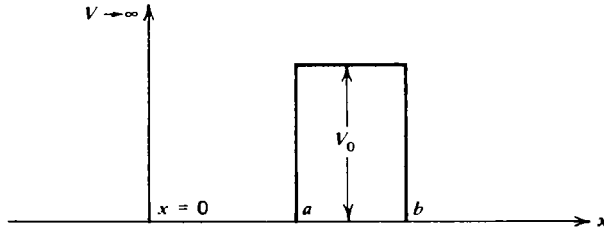
that is, that the matrix has the property

$$S(-k) = S^+(k)$$

4. Without actually solving the Schrödinger equation, set up the solutions so that only the matching of eigenfunctions and their derivatives remain to be done for the following situations:



- (a) flux $\hbar k/m$ would be incident from the left if the potentials were absent; take $E < V_0$.



- (b) with flux of magnitude $\hbar k/m$ incident from the right if the potential were absent; $E < V_0$.

5. Show that the conditions for a bound state (4-51), (4-52) can be obtained by requiring the vanishing of the denominators in (4-26) at $k = i\kappa$. Can you give an argument for why this is not an accident?
6. Consider the scattering matrix for the potential

$$\frac{2m}{\hbar^2} V(x) = \frac{\lambda}{a} \delta(x - b)$$

Show that it has the form

$$\begin{pmatrix} \frac{2ika}{2ika - \lambda} & \frac{\lambda}{2ika - \lambda} e^{-2ikb} \\ \frac{\lambda}{2ika - \lambda} e^{2ikb} & \frac{2ika}{2ika - \lambda} \end{pmatrix}$$

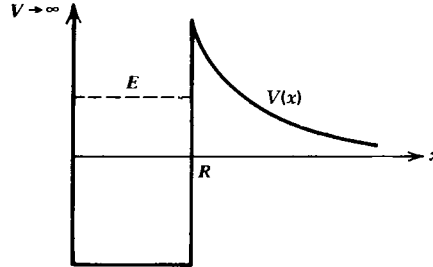
Prove that it is unitary, and that it will yield the condition for bound states when the elements of that matrix become infinite. (This will only occur for $\lambda < 0$.)

7. Consider the harmonic oscillator potential perturbed by a small cubic term, so that

$$V(x) = \frac{1}{2} m\omega^2 \left(x^2 - \frac{x^3}{a} \right)$$

If a is large compared to the characteristic dimension $(\hbar/m\omega)^{1/2}$, the states will all be *metastable*, since there can be no lowest energy state (as $x \rightarrow \infty$, the energy gets arbitrarily negative). Estimate the probability of tunneling from the ground state to the region on the far right.

8. Consider the potential shown in the figure



with

$$V(x) = \frac{\hbar^2 l(l+1)}{2mx^2} \quad x > R$$

Estimate the lifetime of a particle of energy E in this potential. (The outside potential represents a centrifugal barrier in a three-dimensional world.) Express your result in terms of the dimensionless ratio l/kR , where $E = \hbar^2 k^2/2m$, and $l \gg 1$.

9. Consider the following argument: If we have an electron in a potential of width a , then the kinetic energy is, by the uncertainty principle, larger than $\hbar^2/2ma^2$. Thus to get a bound state, the potential energy must not only be negative, but it must also be larger in magnitude than $\hbar^2/2ma^2$. On the other hand, we have shown that in one dimension there is always a bound state, no matter how small V_0 is, provided it is negative. What is wrong with this argument?
10. Consider a potential given by

$$\begin{aligned} V(x) &= \infty & x < 0 \\ &= 0 & x > a \\ &= a \text{ negative function of } x & \text{in between.} \end{aligned}$$

Suppose it is known that the interior wave function is such that

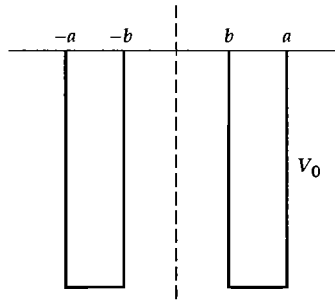
$$\frac{1}{u} \frac{du}{dx} \Big|_{x=a} = f(E)$$

- (a) What is the binding energy of a bound state in terms of $f(E_B)$?
- (b) Suppose $f(E)$ is a very slowly varying function of E , so that we can take it to be a constant. Calculate the reflected amplitude $R(k)$ in terms of f , if the wave function for $x > a$ has the form $e^{-ikx} + R(k)e^{ikx}$, and check that $|R(k)|^2 = 1$.
11. Consider a particle in the double well shown in the figure. Show that the eigenvalue conditions may be written in the form

$$\tan qa - b = \frac{q\alpha(1 + \tanh \alpha b)}{q^2 - \alpha^2 \tanh \alpha b}$$

and

$$\tan q(a - b) = \frac{q\alpha(1 + \coth \alpha b)}{q^2 - \alpha^2 \coth \alpha b}$$



for the even and odd solutions, respectively, where $-E = \hbar^2 \alpha^2 / 2m$ and $E + V_0 = \hbar^2 q^2 / 2m$.

12. Use Figs. 4-7 and 4-8 to discuss the following:
- (a) The condition that there be at most two bound states in the problem
 - (b) The condition that there be at most three bound states in the problem
 - (c) Suppose the third bound state is just barely bound. What can you say about the binding energy of the first and second bound states?
13. Consider the example of Problem 11.
- (a) Show that the eigenvalue conditions approach the single well conditions when $b \rightarrow 0$.
 - (b) Consider the case where the separation between the well centers becomes large, with the well width fixed. Show that the even and odd eigenvalues approach each other in value. Estimate the energy difference between the lowest lying even and odd eigenvalues. (Hint: For large z , $\tanh z = 1 - 2e^{-2z}$. Work to lowest order in e^{-2z} .)
14. Prove the *Virial theorem*, which, in one dimension, takes the form

$$\left\langle \frac{p^2}{2m} \right\rangle = \frac{1}{2} \left\langle x \frac{dV}{dx} \right\rangle$$

To do this (a) show that for real wave functions $\psi(x)$,

$$\int_{-\infty}^{\infty} dx \psi(x) x \frac{dV(x)}{dx} \psi(x) = -\langle V \rangle + 2 \int_{-\infty}^{\infty} dx \frac{d\psi}{dx} x V(x) \psi(x)$$

(b) Use the energy eigenvalue equation to prove that

$$2 \int_{-\infty}^{\infty} dx \frac{d\psi}{dx} x V(x) \psi(x) = E + \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \left(\frac{d\psi}{dx} \right)^2$$

Supplement 4-A

The Wentzel-Kramers-Brillouin-Jeffreys Approximation

This approximation method is particularly useful when one is dealing with slowly varying potentials. Exactly what this means will become clear later. One wants to solve the equation

$$\frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2} [E - V(x)]\psi(x) = 0 \quad (4A-1)$$

and to do so, it is useful to write

$$\psi(x) = R(x)e^{iS(x)/\hbar} \quad (4A-2)$$

Then

$$\frac{d^2\psi}{dx^2} = \left[\frac{d^2R}{dx^2} + \frac{2i}{\hbar} \frac{dR}{dx} \frac{dS}{dx} + \frac{i}{\hbar} R \frac{d^2S}{dx^2} - \frac{1}{\hbar^2} R \left(\frac{dS}{dx} \right)^2 \right] e^{iS(x)/\hbar} \quad (4A-3)$$

so that the differential equation splits into two, by taking the real and imaginary part of (4A-1) after (4A-3) has been substituted. The imaginary part gives

$$R \frac{d^2S}{dx^2} + 2 \frac{dR}{dx} \frac{dS}{dx} = 0 \quad (4A-4)$$

that is,

$$\frac{d}{dx} \left(\log \frac{dS}{dx} + 2 \log R \right) = 0$$

whose solution is

$$\frac{dS}{dx} = \frac{C}{R^2} \quad (4A-5)$$

The real part reads

$$\frac{d^2R}{dx^2} - \frac{1}{\hbar^2} R \left(\frac{dS}{dx} \right)^2 + \frac{2m[E - V(x)]}{\hbar^2} R = 0$$

which, when (4A-5) is substituted, becomes

$$\frac{d^2R}{dx^2} - \frac{C^2}{\hbar^2} \frac{1}{R^3} + \frac{2m[E - V(x)]}{\hbar^2} R = 0 \quad (4A-6)$$

At this point we make the approximation that

$$\frac{1}{R} \frac{d^2R}{dx^2} \ll \frac{C^2}{\hbar^2} \frac{1}{R^4} = \frac{1}{\hbar^2} \left(\frac{dS}{dx} \right)^2 \quad (4A-7)$$

so that the equation becomes

$$\frac{C^2}{R^4} = 2m[E - V(x)] \quad (4A-8)$$

Thus

$$\frac{C}{R^2} = \frac{dS}{dx} = \sqrt{2m[E - V(x)]} \quad (4A-9)$$

and hence

$$S(x) = \int_{x_1}^x dy \sqrt{2m[E - V(x)]} \quad (4A-10)$$

The condition for the validity can be translated into a statement about the variation of $V(x)$. It will be satisfied if $V(x)$ varies slowly in a wavelength, which varies from point to point, but which for slowly varying $V(x)$ is defined by

$$\lambda(x) = \frac{\hbar}{p(x)} = \frac{\hbar}{[2m[E - V(x)]]^{1/2}} \quad (4A-11)$$

At the points where

$$E - V(x) = 0 \quad (4A-12)$$

special treatment is required, because in the approximate equation (4A-8) $R(x)$ appears to be singular. This cannot be, which means that the approximation (4A-7) must be poor there. The special points are called *turning points* because it is there that a classical particle would turn around: It can only move where $E - V(x) \geq 0$. The way of handling solutions near turning points is a little too technical to be presented here. The basic idea is that we have a solution to the left of the turning point [where $E > V(x)$, say], of the form

$$\psi(x) = R(x) \exp \left\{ i \int_{x_1}^x dy \sqrt{2m(E - V(y))/\hbar^2} \right\} \quad (4A-13)$$

and a solution to the right of the turning point [where $E < V(x)$], and what we need is a formula that interpolates between them. In the vicinity of the turning point one can approximate $\sqrt{(2m/\hbar^2)[E - V(x)]}$ by a straight line over a small interval, and solve the Schrödinger equation exactly. Since it is a second-order equation, there are two adjustable constants, one that is fixed by fitting the solution to (4A-13) and the other by fitting it to

$$\psi(x) = R(x) \exp \left\{ - \int_{x_1}^x dy \sqrt{2m(V(y) - E)/\hbar^2} \right\} \quad (4A-14)$$

the solution to the right of the turning point.¹ The preceding solution decreases in amplitude as x increases. The total attenuation at the next turning point, when $E \geq V(x)$ again, is

$$\frac{\psi(x_{II})}{\psi(x_I)} \approx \exp \left\{ - \int_{x_I}^{x_{II}} dy \sqrt{2m(V(y) - E)/\hbar^2} \right\} \quad (4A-15)$$

which is just the square root of the transmission probability that we found in Chapter 4.

¹For more details, see almost any of the more advanced books on quantum mechanics—for example, J. L. Powell and B. Crasemann, *Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1961; L. I. Schiff, *Quantum Mechanics*, McGraw-Hill, New York, 1968.

Supplement 4-B

Tunneling in Nuclear Physics

Tunneling is important in nuclear physics. Nuclei are very complicated objects, but under certain circumstances it is appropriate to view nucleons as independent particles occupying levels in a potential well. With this picture in mind, the decay of a nucleus into an α -particle (a He nucleus with $Z = 2$) and a daughter nucleus can be described as the tunneling of an α -particle through a barrier caused by the Coulomb potential between the daughter and the α -particle (Fig. 4B-1). The α -particle is not viewed as being in a bound state: if it were, the nucleus could not decay. Rather, the α -particle is taken to have positive energy, and its escape is only inhibited by the existence of the barrier.¹

If we write

$$|T|^2 = e^{-G} \quad (4B-1)$$

then

$$G = 2\left(\frac{2m}{\hbar^2}\right)^{1/2} \int_R^b dr \sqrt{\frac{Z_1 Z_2 e^2}{4\pi\epsilon_0 r} - E} \quad (4B-2)$$

where R is the nuclear radius² and b is the turning point, determined by the vanishing of the integrand (4B-2); Z_1 is the charge of the daughter nucleus, and $Z_2 (= 2 \text{ here})$ is the charge of the particle being emitted. The integral can be done exactly

$$\int_R^b dr \left(\frac{1}{r} - \frac{1}{b}\right)^{1/2} = \sqrt{b} \left[\cos^{-1}\left(\frac{R}{b}\right)^{1/2} - \left(\frac{R}{b} - \frac{R^2}{b^2}\right)^{1/2} \right] \quad (4B-3)$$

At low energies (relative to the height of the Coulomb barrier at $r = R$), we have $b \gg R$, and then

$$G = \frac{2}{\hbar} \left(\frac{2mZ_1 Z_2 e^2 b}{4\pi\epsilon_0} \right)^{1/2} \left[\frac{\pi}{2} \sqrt{\frac{R}{b}} \right] \quad (4B-4)$$

with $b = Z_1 Z_2 e^2 / 4\pi\epsilon_0 E$. If we write for the α -particle energy $E = mv^2/2$, where v is its final velocity, then

$$G = \frac{2\pi Z_1 Z_2 e^2}{4\pi\epsilon_0 \hbar v} = 2\pi\alpha Z_1 Z_2 \left(\frac{c}{v}\right) \quad (4B-5)$$

¹If you find it difficult to imagine why a repulsion would keep two objects from separating, think of the inverse process, α capture. It is clear that the barrier will tend to keep the α -particle out.

²In fact, early estimations of the nuclear radius came from the study of α -decay. Nowadays one uses the size of the charge distribution as measured by scattering electrons off nuclei to get nuclear radii. It is not clear that the two should be expected to give exactly the same answer.

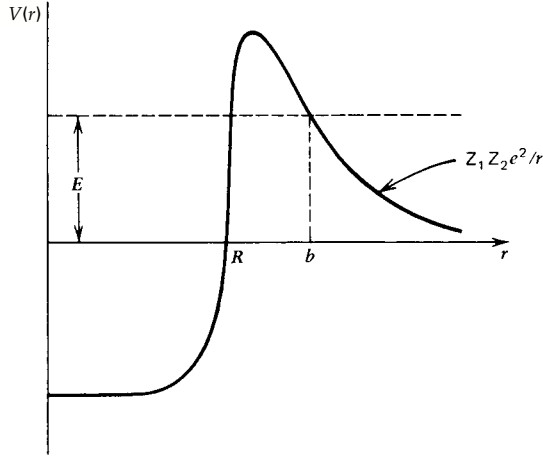


Figure 4B-1 Potential barrier for α decay.

The time taken for an α -particle to get out of the nucleus may be estimated as follows: the probability of getting through the barrier on a single encounter is e^{-G} . Thus the number of encounters needed to get through is $n \simeq e^G$. The time between encounters is of the order of $2R/v$, where R is again the nuclear radius, and v is the α velocity inside the nucleus. Thus the lifetime is

$$\tau \simeq \frac{2R}{v} e^G \quad (4B-6)$$

The velocity of the α inside the nucleus is a rather fuzzy concept, and the whole picture is very classical, so that the factor in front of the e^G cannot really be predicted without a much more adequate theory. Our considerations do give us an order of magnitude for it. For a 1-MeV α -particle,

$$v = \sqrt{\frac{2E}{m}} = c \sqrt{\frac{2E}{mc^2}} = 3 \times 10^8 \sqrt{\frac{2}{4 \times 940}} \simeq 7.0 \times 10^6 \text{ m/s}$$

so that one predicts, for low energy α 's, the straight-line plot

$$\log_{10} \frac{1}{\tau} \simeq \text{const} - 1.73 \frac{Z_1}{\sqrt{E(\text{MeV})}} \quad (4B-7)$$

with the constant in front of the order of magnitude 27–28 when τ is measured in years instead of seconds. A large collection of data shows that a good fit to the lifetime data is obtained with the formula

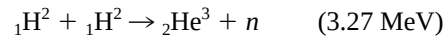
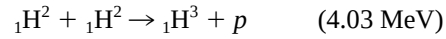
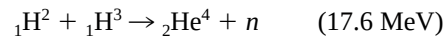
$$\log_{10} \frac{1}{\tau} = C_2 - C_1 \frac{Z_1}{\sqrt{E}} \quad (4B-8)$$

Here $C_1 = 1.61$ and C_2 lying between 55 and 62. The exponential part of the fit differs slightly from our derivation, but given the simplicity of our model, the agreement has to be rated as good.

For larger α energies, the G factor depends on R , and with $R = r_0 A^{1/3}$, one finds that r_0 is a constant—that is, that the notion of a Coulomb barrier taking over the role of the potential beyond the nuclear radius has some validity. Again, simple qualitative considerations explain the data.

W-18 Supplement 4-B Tunneling in Nuclear Physics

The fact that the probability of a reaction (e.g., capture) between nuclei is attenuated by the factor $e^{-2(Z_1Z_2/\sqrt{E})}$ ($Z_2 = 2$ for α 's) implies that at low energies and/or for high Z 's such reactions are rare. That is why all attempts to make thermonuclear reactors concentrate on the burning of hydrogen (actually heavy hydrogen–deuterium).



since reactions involving higher Z elements would require much higher energies, that is, much higher temperatures, with correspondingly greater confinement problems. For the same reason, neutrons are used in nuclear reactors to fission the heavy elements. Protons, at the low energies available, would not be able to get near enough to the nuclei to react with them.

Supplement 4-C

Periodic Potentials

Metals generally have a crystalline structure; that is, the ions are arranged in a way that exhibits a spatial periodicity. In our one-dimensional discussion of this topic, we will see that this periodicity has two effects on the motion of the free electrons in the metal. One is that for a perfect *lattice*—that is, for ions spaced equally—the *electron propagates without reflection*; the other is that there are restrictions on the energies allowed for the electrons; that is, *there are allowed and forbidden energy “bands.”*

We begin with a discussion of the consequences of perfect periodicity.

The periodicity will be built into the potential, for which we require that

$$V(x + a) = V(x) \quad (4C-1)$$

Since the kinetic energy term $-(\hbar^2/2m)(d^2/dx^2)$ is unaltered by the change $x \rightarrow x + a$, the whole *Hamiltonian is invariant under displacements by a* . For the case of zero potential, when the solution corresponding to a given energy $E = \hbar^2 k^2/2m$ is

$$\psi(x) = e^{ikx} \quad (4C-2)$$

the displacement yields

$$\psi(x + a) = e^{ik(x+a)} = e^{ika} \psi(x) \quad (4C-3)$$

that is, the original solution multiplied by a phase factor, so that

$$|\psi(x + a)|^2 = |\psi(x)|^2 \quad (4C-4)$$

The observables will therefore be the same at x as at $x + a$; that is, we cannot tell whether we are at x or at $x + a$. In our example we shall also insist that $\psi(x)$ and $\psi(x + a)$ differ only by a phase factor, which need not, however, be of the form e^{ika} .

We digress briefly to discuss this requirement more formally. The invariance of the Hamiltonian under a displacement $x \rightarrow x + a$ can be treated formally as follows. Let D_a be an operator whose rule of operation is that

$$D_a f(x) = f(x + a) \quad (4C-5)$$

The invariance implies that

$$[H, D_a] = 0 \quad (4C-6)$$

We can find the eigenvalues of this operator by noting that

$$D_a \psi(x) = \lambda_a \psi(x) \quad (4C-7)$$

together with

$$D_{-a} D_a f(x) = D_a D_{-a} f(x) = f(x) \quad (4C-8)$$

implies that $\lambda_a \lambda_{-a} = 1$. This then implies that λ_a must be of the form e^{iqa} . Here q must be real, because if q had an imaginary part, a succession of displacements by a would make the wave function larger and larger with each displacement in one or the other direction.

Consider now a simultaneous eigenfunction of H and D_a , and define

$$u(x) = e^{-iqx}\psi(x) \quad (4C-9)$$

Then, using the fact that $\lambda_a = e^{iqa}$, we get

$$D_a u(x) = e^{-iq(x+a)} D_a \psi(x) = e^{-iq(x+a)} e^{iqa} \psi(x) = e^{-iqx} \psi(x) = u(x) \quad (4C-10)$$

This means that $u(x)$ is a periodic function obeying $u(x+a) = u(x)$. The upshot is that a function which is a simultaneous eigenfunction of H and D_a must be of the form

$$\psi(x) = e^{iqx} u(x) \quad (4C-11)$$

with $u(x)$ periodic. This result is known as *Bloch's Theorem*.

For a free particle $q = k$, the wave number corresponds to the energy E . More generally, the relation between q and k is more complicated. In any case, it is clear that (4C-4) holds, so that the net flux is unchanged as we go from x to $x+a$, and by extension to $x+na$. This means that an electron propagates without a change in flux.

Let us consider a series of ions in a line, with their centers located at $x = na$. To avoid having to deal with *end effects*, we assume that there are N ions placed on a very large ring, so that $n = 1$ and $n = N+1$ are the same site. We will assume that the most loosely bound electrons—the ones that are viewed as “free”—are still sufficiently strongly bound to the ions that their wave functions do not overlap more than one or two nearest neighbors. We may now ask: What is the effect of this overlap on the energies of the electrons?

To answer this question, we consider first a classical analogy. We represent the electrons at the different sites by simple harmonic oscillators, all oscillating with the same angular frequency ω . In the absence of any coupling between the oscillators, we have the equation of motion

$$\frac{d^2 x_n}{dt^2} = -\omega^2 x_n \quad (n = 0, 1, 2, \dots) \quad (4C-12)$$

If the harmonic oscillators are coupled to their nearest neighbors, then the equation is changed to

$$\frac{d^2 x_n}{dt^2} = -\omega^2 x_n - K[(x_n - x_{n-1}) + (x_n - x_{n+1})] \quad (4C-13)$$

To solve this we write down a trial solution

$$x_n = A_n \cos \Omega t \quad (4C-14)$$

When this is substituted into (4C-13), we get

$$(\omega^2 - \Omega^2)A_n = -K(2A_n - A_{n-1} - A_{n+1}) \quad (4C-15)$$

This is known as a *difference equation*. We solve it by a trial solution. Let us assume that

$$A_n = L^n \quad (4C-16)$$

The identification of the sites at $n = 1$ and $N+1$ implies $A_1 = A_{N+1}$ so that $L^N = 1$. This means that

$$L = e^{2\pi ir/N} \quad r = 0, 1, 2, \dots, (N-1) \quad (4C-17)$$

The equation for the frequency Ω now yields

$$\omega^2 - \Omega^2 = -2K \left(1 - \cos \frac{2\pi r}{N} \right) = -4K \sin^2 \frac{\pi r}{N}$$

The result

$$\Omega^2 = \omega^2 + 4K \sin^2 \frac{\pi r}{N} \quad (4C-18)$$

shows that the frequencies, which, without coupling are all ω —that is, are *N-fold degenerate* (which corresponds to all the pendulums moving together)—are now spread over a range from ω to $\sqrt{\omega^2 + 4K}$. For large N there are many such frequencies, and they can be said to form a *band*. If we think of electrons as undergoing harmonic oscillations about their central locations, we can translate the above into a statement that in the absence of neighbors, all electron energies are degenerate, and the interaction with neighboring atoms spreads the energy values. We can, of course, have several fundamental frequencies $\omega_1, \omega_2, \dots$, and different couplings to their neighbors, with strengths $K_1, K_2, \dots$, which will then give rise to several bands that may or may not overlap.

The spreading of the frequencies is the same effect as the spreading of the energy levels of the most loosely bound electrons. For atoms far apart, with spacing larger than the exponential fall-off of the wave functions, all the energies are the same so that we have an *N-fold degenerate* single energy. Because the atoms are not so far apart, there is some coupling between nearest neighbors, and the energy levels spread. The classical analogy is suggestive, but not exact, since for the quantum case levels are pushed up as well as down, whereas all the frequencies above, lie above ω . Later we solve the Kronig-Penney model in which the potential takes the form

$$V(x) = \frac{\hbar^2 \lambda}{2m a} \sum_{-\infty}^{\infty} \delta(x - na) \quad (4C-19)$$

The solution can be shown to lead to a condition on q , which reads

$$\cos qa = \cos ka + \frac{1}{2} \lambda \frac{\sin ka}{ka} \quad (4C-20)$$

As can be seen from Figure (4C-1), this clearly shows the energy band structure.

THE KRONIG-PENNEY MODEL

To simplify the algebra, we will take a series of repulsive delta-function potentials,

$$V(x) = \frac{\hbar^2 \lambda}{2m a} \sum_{n=-\infty}^{\infty} \delta(x - na) \quad (4C-21)$$

Away from the points $x = na$, the solution will be that of the free-particle equation—that is, some linear combination of $\sin kx$ and $\cos kx$ (we deal with real functions for simplicity). Let us assume that in the region R_n defined by $(n - 1)a \leq x \leq na$, we have

$$\psi(x) = A_n \sin k(x - na) + B_n \cos k(x - na) \quad (4C-22)$$

and in the region R_{n+1} defined by $na \leq x \leq (n + 1)a$ we have

$$\psi(x) = A_{n+1} \sin k[x - (n + 1)a] + B_{n+1} \cos k[x - (n + 1)a] \quad (4C-23)$$

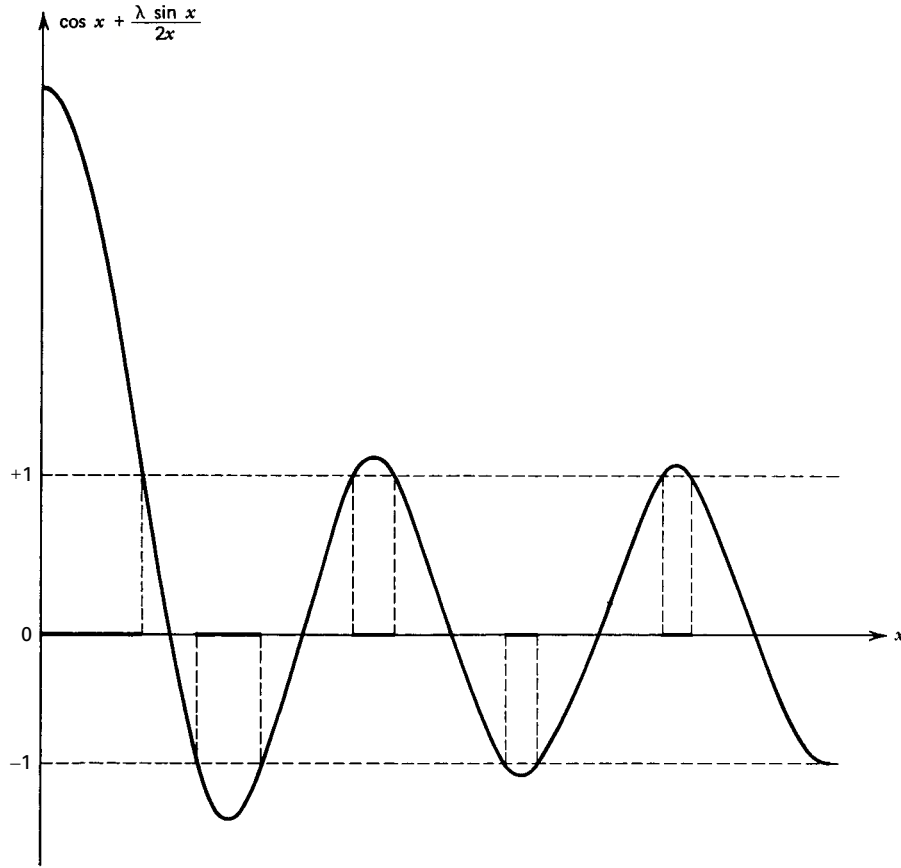


Figure 4C-1 Plot of $\cos x + (\lambda/2)(\sin x/x)$ as a function of x . The horizontal lines represent the bounds ± 1 . The regions of x for which the curve lines outside the strip are forbidden.

Continuity of the wave function implies that ($x = na$)

$$-A_{n+1} \sin ka + B_{n+1} \cos ka = B_n \quad (4C-24)$$

and the discontinuity condition (4-68) here reads

$$kA_{n+1} \cos ka + kB_{n+1} \sin ka - kA_n = \frac{\lambda}{a} B_n \quad (4C-25)$$

A little manipulation yields

$$\begin{aligned} A_{n+1} &= A_n \cos ka + (g \cos ka - \sin ka) B_n \\ B_{n+1} &= (g \sin ka + \cos ka) B_n + A_n \sin ka \end{aligned} \quad (4C-26)$$

where $g = \lambda/ka$.

The requirement from Bloch's theorem that

$$\psi(x+a) = e^{iq(x+a)}u(x+a) = e^{iq(x+a)}u(x) = e^{iqa}\psi(x) \quad (4C-27)$$

implies that the wave functions in the adjacent regions R_n and R_{n+1} are related, since the wave function in (4C-22) may be written as

$$\psi(x) = A_n \sin[k((x+a)-(n+1)a)] + B_n \cos k[(x+a)-(n+1)a]$$

which is identical to that in (4C-23), provided

$$\begin{aligned} A_{n+1} &= e^{iqa} A_n \\ B_{n+1} &= e^{iqa} B_n \end{aligned} \quad (4C-28)$$

When this is inserted into the (4C-26), that is, into the conditions that the wave equation obeys the Schrödinger equation with the delta function potential, we get

$$\begin{aligned} A_n(e^{iqa} - \cos ka) &= B_n(g \cos ka - \sin ka) \\ B_n(e^{iqa} - (g \sin ka + \cos ka)) &= A_n \sin ka \end{aligned} \quad (4C-29)$$

This leads to the condition

$$(e^{iqa} - \cos ka)(e^{iqa} - (g \sin ka + \cos ka)) = \sin ka(g \cos ka - \sin ka) \quad (4C-30)$$

This may be rewritten in the form

$$e^{2iqa} - 2(\cos ka + \frac{g}{2} \sin ka)e^{iqa} + 1 = 0 \quad (4C-31)$$

This quadratic equation can be solved, and both real and imaginary parts lead to the condition

$$\cos qa = \cos ka + \frac{\lambda}{2} \frac{\sin ka}{ka} \quad (4C-32)$$

This is a very interesting result, because the left side is always bounded by 1; that is, there are restrictions on the possible ranges of the energy $E = \hbar^2 k^2 / 2m$ that depend on the parameters of our “crystal.” Figure 4C-1 shows a plot of the function $\cos x + \lambda \sin x / 2x$ as a function of $x = ka$. The horizontal line represents the bounds on $\cos qa$, and the regions of x , for which the curve lies outside the strip, are forbidden regions. Thus there are *allowed energy bands* separated by regions that are forbidden. Note that the onset of a forbidden band corresponds to the condition

$$qa = n\pi \quad n = \pm 1, \pm 2, \pm 3, \dots \quad (4C-33)$$

This, however, is just the condition for Bragg reflection with normal incidence. The existence of energy gaps can be understood qualitatively. In first approximation the electrons are free, except that there will be Bragg reflection when the waves reflected from successive atoms differ in phase by an integral number of 2π —that is, when (4C-33) is satisfied. These reflections give rise to standing waves, with even and odd waves of the form $\cos \pi x/a$ and $\sin \pi x/a$, respectively. The energy levels corresponding to these standing waves are degenerate. Once the attractive interaction between the electrons and the positively charged ions at $x = ma$ (m integer) is taken into account, the even states, peaked in between, will move up in energy. Thus the energy degeneracy is split at $q = n\pi/a$, and this leads to energy gaps, as shown in Fig. 4C-1.

The Kronig-Penney model has some relevance to the theory of metals, insulators, and semiconductors if we take into account the fact (to be studied later) that energy levels

occupied by electrons cannot accept more electrons. Thus a metal may have an energy band partially filled. If an external field is applied, the electrons are accelerated, and if there are momentum states available to them, the electrons will occupy the momentum states under the influence of the electric field. Insulators have completely filled bands, and an electric field cannot accelerate electrons, since there are no neighboring empty states. If the electric field is strong enough, the electrons can “jump” across a forbidden energy gap and go into an empty allowed energy band. This corresponds to the breakdown of an insulator. The semiconductor is an insulator with a very narrow forbidden gap. There, small changes of conditions, such as a rise in temperature, can produce the “jump” and the insulator becomes a conductor.

The band structure is of great relevance in solid state physics. Fig. 4C-2 shows three situations that can arise when energy levels are filled with electrons. We shall learn in Chapter 13 that only two electrons are allowed per energy level. In case (a) the electrons fill all the energy levels below the edge of the energy gap. The application of a weak electric field will have no effect on the material. The electrons near the top of the filled band cannot be accelerated. There are no levels with higher energy available to them. Materials in which this occurs are *insulators*; that is, they do not carry currents when electric fields are applied. In case (b) the energy levels are only partly filled. In this case the application of an electric field accelerates the electrons at the top of the stack of levels. These electrons have empty energy levels to move into, and they would accelerate indefinitely in a perfect lattice, as stated in the previous section. What keeps them from doing that is *dissipation*. The lattice is not perfect for two reasons: one is the presence of impurities, which destroys the perfect periodicity; the other is the effect of thermal agitation on the position of the ions forming the lattice, which has the same effect of destroying perfect periodicity. Materials in which the energy levels below the gaps are only partially filled are *conductors*.

The width of the gaps in the energy spectrum depends on the materials. For some insulators the gaps are quite narrow. When this happens, then at finite temperatures T , there is a calculable probability that some of the electrons are excited to the bottom of the set of energy levels above the gap. (To good approximation the probability is proportional to the Boltzmann factor $e^{-E/kT}$). These electrons can be accelerated as in a conductor, so that the application of an electric field will give rise to a current. The current is augmented by another effect: the energy levels that had been occupied by the electrons promoted to the higher energy band (called the *conduction band*) are now empty. They provide vacancies into which electrons in the lower band (called the *valence band*) can be accelerated into,

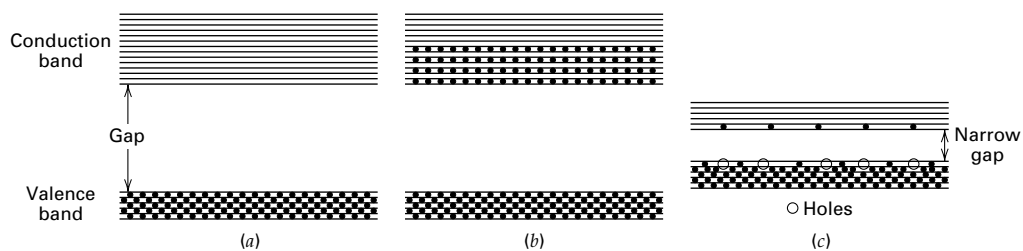


Figure 4C-2 Occupation of levels in the lowest two energy bands, separated by a gap. (a) Insulator has a completely filled band. Electrons cannot be accelerated into a nearby energy level. (b) Conductor has a half-filled band, allowing electrons to be accelerated into nearby energy levels. (c) In a semiconductor, thermal effects promote some electrons into a second band. These electrons can conduct electricity. The electrons leave behind them *holes* that act as positively charged particles and also conduct electricity.

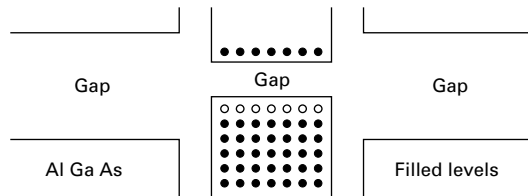


Figure 4C-3 Schematic picture of electrons and holes trapped in a well created by adjacent semiconductors with a wider gap. An example of such a heterostructure is provided by a layer of GaAs sandwiched between two layers of AlGaAs.

when an electric field is applied. These vacancies, called *holes*, propagate in the direction opposite to that of the electrons and thus add to the electric current. This is the situation shown in Fig. 4C-2(c).

The technology of making very thin layers of compounds of materials has improved in recent decades to such an extent that it is possible to create the analog of the infinite wells discussed in Chapter 3. Consider a “sandwich” created by two materials. The outer one has a larger energy gap than the inner one, as shown in Fig. 4C-3. The midpoints of the gaps must coincide¹ (for equilibrium reasons). The result is that both electrons and holes in the interior semiconductor cannot move out of the region between the outer semiconductors, because there are no energy levels that they can move to. Such confined regions may occur in one, two, or three dimensions. In the last case we deal with *quantum dots*. The study of the behavior of electrons in such confined regions is a very active field of research in the study of materials.

In summary, one-dimensional problems give us a very important glimpse into the physics of quantum systems in the real world of three dimensions.

¹A brief, semiquantitative discussion of this material may be found in *Modern Physics* by J. Bernstein, P. M. Fishbane, and S. Gasiorowicz (Prentice Hall, 2000). See also Chapter 44 in *Physics for Scientists and Engineers*, (2nd Edition) by P. M. Fishbane, S. Gasiorowicz and S. T. Thornton (Prentice Hall, 1996). There are, of course, many textbooks on semiconductors, which discuss the many devices that use *bandgap engineering* in great quantitative detail. See in particular L. Solymar and D. Walsh, *Lectures on the Electrical Properties of Materials*, Oxford University Press, New York (1998).

Chapter 5

The General Structure of Wave Mechanics

In this chapter we recapitulate what we learned in Chapter 3, and generalize from our specific examples. This involves learning more about operators, hermitian operators, the expansion theorem, and degeneracy. We also go beyond our earlier discussion in treating the time dependence of expectation values of operators.

5-1 EIGENFUNCTIONS AND EIGENVALUES

We assert that *the state of a physical system is described by a wave function, and that this wave function contains all the information that we can have about the system*. This statement needs elaboration. We begin by discussing the time evolution of a physical system. For notational simplicity we shall deal with the wave function of a single particle moving in one dimension, though some of the things that we will discuss don't really appear till we deal with three dimensional systems. For the present we deal with the wave function $\psi(x, t)$.

The Hamiltonian Operator

The time dependence of the wave function is given by

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = H\psi(x, t) \quad (5-1)$$

On the right side of the equation the wave function is acted upon by an *operator* H , the Hamiltonian, which, because it governs the time development of the system, plays a central role in quantum mechanics. For a single particle moving in a potential $V(x)$, it is an operator version of the total energy

$$H = \frac{p_{\text{op}}^2}{2m} + V(x) \quad (5-2)$$

where the momentum is represented by its operator version

$$p_{\text{op}} = -i\hbar \frac{\partial}{\partial x} \quad (5-3)$$

If $V(x)$ has no explicit time dependence, then (5-1) can be partially solved by writing

$$\psi(x, t) = u_E(x)e^{-iEt/\hbar} \quad (5-4)$$

The function $u_E(x)$ is a solution of the differential equation

$$Hu_E(x) = Eu_E(x) \quad (5-5)$$

The solution has the characteristic that H acting on it reproduces it with a multiplicative constant E . Such solutions are called *eigensolutions*, and the constants E that satisfy the equation together with the necessary *boundary conditions* are called *eigenvalues*.

Some properties of eigenvalues and eigenfunctions are familiar from our examples in Chapter 3.

1. Eigenfunctions that correspond to different eigenvalues are *orthogonal*; that is,

$$\int_{-\infty}^{\infty} dx u_{E_1}^*(x) u_{E_2}(x) = 0 \quad (5-6)$$

when $E_1 \neq E_2$.

2. The eigenfunctions of H form a complete set. By this we mean the following: Any *arbitrary* square integrable function of x , one that satisfies

$$\int_{-\infty}^{\infty} dx |\psi(x)|^2 < \infty$$

may be expanded in terms of eigenfunctions of H , so that

$$\psi(x) = \sum_E C_E u_E(x) \quad (5-7)$$

where the sum is over *all* the eigenfunctions. The eigenvalues of H , the *spectrum of H* , may be a discrete set as in the infinite box, or it may involve discrete values (labeled by n as for the well) and also as a continuous energy variable, as appears in the example of a plane wave traversing a potential well seen in Chapter 4. In that latter case we have the schematic expression

$$\psi(x) = \sum_n C_n u_n(x) + \int dE C(E) u_E(x) \quad (5-8)$$

The expression is schematic because we have not specified exactly which form of the continuum eigenfunctions we are supposed to use. Recall that for a square well, our positive energy solutions were chosen such that they involved an incoming plane wave from $x = -\infty$. We could equally well have taken a solution that involves an incoming plane wave from $x = +\infty$, or any linear combination of the two. We will not dwell on this at this point.

3. The eigenfunctions can be multiplied by constants so that they become *normalized*. For the discrete set this reads

$$\int_{-\infty}^{\infty} dx u_m^*(x) u_n(x) = \delta_{mn} \quad (5-9)$$

4. Consider any solution of (5-1), $\psi(x, t)$. At $t = 0$, the wave function $\psi(x, 0)$, which we write as $\psi(x)$, may be expanded as in (5-7), for example. Because each eigenfunction has a simple time dependence given by $e^{-iEt/\hbar}$, the time-dependent solution has the form

$$\psi(x, t) = \sum_E C_E u_E(x) e^{-iEt/\hbar} \quad (5-10)$$

We already saw the more general form, corresponding to the expansion (5-8), in (3-12).

5-2 OTHER OBSERVABLES

The energy is just one observable attribute of a system. We discussed other observables such as momentum, and we will discuss angular momentum in a later chapter. We recall that for the momentum, the eigenfunctions are solutions of the equation

$$p_{\text{op}} u_p(x) = -i\hbar \frac{du_p(x)}{dx} = p u_p(x) \quad (5-11)$$

The eigenvalues form a continuum ($-\infty < p < \infty$), and the eigenfunctions are

$$u_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \quad (5-12)$$

with the factor in front chosen such that the orthonormality relation reads

$$\int_{-\infty}^{\infty} dx u_{p_1}^*(x) u_{p_2}(x) = \delta(p_1 - p_2) \quad (5-13)$$

The expansion theorem in this case is usually written in the form

$$\psi(x) = \int_{-\infty}^{\infty} dp \phi(p) u_p(x) \quad (5-14)$$

Note that what we called “the wave function in momentum space” is really just an expansion coefficient, and with our interpretation, its absolute square $|\phi(p)|^2 dp$ gives us the probability that a momentum measurement on the system described by $\psi(x)$ yields a value lying in the range $(p, p + dp)$.

The momentum operator, like the Hamiltonian, has *real* eigenvalues, which is what we expect of all operators that represent physical quantities. Operators that have all their eigenvalues real are called *hermitian* operators. We can make a general statement:

Since all physical observables must have real eigenvalues, they must be represented by hermitian operators.

The completeness property of eigenfunctions is characteristic of hermitian operators, and we may therefore generalize our statements as follows:

For an arbitrary observable, which we denote by A , there will be eigenfunctions corresponding to the real eigenvalues a ,

$$A u_a(x) = a u_a(x) \quad (5-15)$$

The eigenfunctions form an orthogonal set, and they can be normalized so that

$$\int_{-\infty}^{\infty} dx u_{a_1}^*(x) u_{a_2}(x) = \delta_{a_1 a_2} \quad (5-16)$$

Here $\delta_{a_1 a_2}$ is a Kronecker delta for discrete eigenvalues and a Dirac delta function for continuous eigenvalues.

The eigenfunctions $u_a(x)$ also form a complete set, which is equivalent to the statement that an arbitrary (square integrable) function $\psi(x)$ may be expanded in terms of the $u_a(x)$,

$$\psi(x) = \sum_a C_a u_a(x) \quad (5-17)$$

It follows from the orthonormality condition that

$$C_a = \int_{-\infty}^{\infty} dx u_a^*(x) \psi(x) \quad (5-18)$$

The Interpretation of the Expansion Coefficients

The interpretation of the expansion coefficients C_a is the following: If the observable A is measured for a collection of systems each of which is described by the wave function $\psi(x)$, which is normalized to unity, so that

$$\int_{-\infty}^{\infty} dx \psi^*(x) \psi(x) = 1$$

then

1. The result of any given measurement can only be one of the eigenvalues a .
2. The probability that the eigenvalue a will be found, or, equivalently, the fraction of systems in the collection that will be found to have the eigenvalue a , is $|C_a|^2$.
3. After a measurement on a member of the collection yields a given eigenvalue a_1 , for example, then that particular system in the collection must be *projected* by the measurement into the state $u_{a_1}(x)$. Only in this way can we be sure that an immediate repetition of the measurement of the observable A gives the same result.

A consequence of this interpretation is that the probability that the value of the observable A for a system has any one of the eigenvalues is unity. This means that

$$\sum_a |C_a|^2 = 1 \quad (5-19)$$

which follows from

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} dx \psi^*(x) \psi(x) = \int_{-\infty}^{\infty} dx \left(\sum_a C_a^* u_a^*(x) \right) \psi(x) \\ &= \sum_a C_a^* C_a \end{aligned}$$

This formula implies the following:

$$\begin{aligned} \sum_a C_a^* C_a &= \sum_a \int_{-\infty}^{\infty} dx u_a^*(x) \psi(x) \int_{-\infty}^{\infty} dy u_a(y) \psi^*(y) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy \psi^*(y) \psi(x) \sum_a u_a(y) u_a^*(x) = 1 \end{aligned}$$

which, in turn, implies

$$\sum_a u_a(y) u_a^*(x) = \delta(x - y) \quad (5-20)$$

This property of the eigenfunctions is described as the *completeness* relation, and it is equivalent to the statement of the expansion theorem.

We may use the completeness statement to obtain some useful relations. As a simple example, consider the normalized eigenfunctions for the infinite well

$$u_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi x}{a}$$

Equation (5-20) implies that

$$\sum_{n=1}^{\infty} \frac{2}{a} \sin \frac{n\pi x}{a} \sin \frac{n\pi y}{a} = \delta(x - y)$$

Using

$$a\delta(x - y) = \delta\left(\frac{x}{a} - \frac{y}{a}\right)$$

and changing notation so that $x/a = X$ and $y/a = Y$, we end up with

$$\sum_{n=-\infty}^{\infty} \sin n\pi X \sin n\pi Y = \delta(X - Y) \quad (5-21)$$

Another example uses the momentum eigenfunctions

$$u_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$$

In (5-20) we now replace the sum by an integral, since the variable p is continuous. We thus get

$$\delta(x - y) = \int_{-\infty}^{\infty} dp u_p^*(x) u_p(y) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp e^{ip(y-x)/\hbar} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(y-x)} \quad (5-22)$$

We found this result in Supplement 2-A [www.wiley.com/college/gasiorowicz].

5-3 VECTOR SPACES AND OPERATORS

In Chapter 3 we drew attention to the fact that the complete set of eigenfunctions was similar to a set of *unit vectors* in a vector space, and the expansion theorem was similar to the expansion of an arbitrary vector in terms of unit vectors that span the vector space (see p. 52). There is more to this similarity. In Chapter 6 we shall develop the idea that the set of square integrable functions $\psi(x)$ are just realizations of *states* that are represented by vectors in some abstract linear space. The vectors in this space obey familiar rules. For example, the sum of two vectors is a vector—which translates into the statement that the sum of two wave functions is a wave function. There is a scalar product (generalization of the *dot product*) in the vector space, and the realization of this is an integral of the form

$$\int_{-\infty}^{\infty} dx \phi^*(x) \psi(x) \quad (5-23)$$

The vector space is generally complex, and continuously infinite-dimensional since x , the “label” on $\psi(x)$, is continuous. This leads to mathematical complication in that questions of convergence of integrals need to be dealt with. The mathematical theory in which such matters are studied is the theory of *Hilbert spaces*. This level of mathematical sophistication is not needed for our purposes, and we shall for the most part ignore the fact that the

states are vectors in a Hilbert space, and pretend that aside from the fact that the vectors are complex, we can lean on analogies with finite vector spaces.

In vector spaces an operator transforms one vector into another. Linear operators (the only ones of interest to us) have the property that

$$A(\alpha_1\psi_1(x) + \alpha_2\psi_2(x)) = \alpha_1A\psi_1(x) + \alpha_2A\psi_2(x) \quad (5-24)$$

where A is the operator and α_1 and α_2 are arbitrary complex numbers. For any linear operator Q , we may define a hermitian conjugate operator, denoted by Q^+ , by the relation

$$\int_{-\infty}^{\infty} dx \psi^*(x) Q^+ \psi(x) = \int_{-\infty}^{\infty} dx (Q\psi(x))^* \psi(x) \quad (5-25)$$

for any permissible (that is, square integrable) wave function. Linear operators that represent observables have the special property that their expectation value for all of the states that form the linear space must be real. Such operators are called *hermitian* or sometimes *self-adjoint*. The reality of the expectation values

$$\langle A \rangle^* = \langle A \rangle \quad (5-26)$$

translates into the statement that for any wave function $\psi(x)$,

$$\int_{-\infty}^{\infty} dx \psi^*(x) A \psi(x) = \left[\int_{-\infty}^{\infty} dx \psi^*(x) A \psi(x) \right]^* = \int_{-\infty}^{\infty} dx (A \psi(x))^* \psi(x) \quad (5-27)$$

Comparison with (5-25) shows that a hermitian operator is one that is equal to its conjugate; that is,

$$A = A^+ \quad (5-28)$$

1. A consequence of (5-27) is that for hermitian operators,

$$\int_{-\infty}^{\infty} dx \phi^*(x) A \psi(x) = \int_{-\infty}^{\infty} dx (A \phi(x))^* \psi(x)$$

for any admissible pair of wave functions $\phi(x)$ and $\psi(x)$. We leave the proof of this statement to the reader (see Problem 5-1).

2. *Eigenfunctions of a hermitian operator corresponding to different eigenvalues are orthogonal.* This means that if

$$A u_1(x) = a_1 u_1(x) \quad (5-29a)$$

and

$$A u_2(x) = a_2 u_2(x) \quad (5-29b)$$

with $a_1 \neq a_2$, then¹

$$\int_{-\infty}^{\infty} dx u_1^*(x) u_2(x) = 0 \quad (5-30)$$

¹We will prove this result in the next chapter as an exercise in the use of the powerful *Dirac notation*

3. For any two operators A and B ,

$$\int_{-\infty}^{\infty} dx (AB\phi(x))^* \psi(x) = \int_{-\infty}^{\infty} dx \phi^*(x) (AB)^+ \psi(x) = \int_{-\infty}^{\infty} dx \phi^*(x) B^+ A^+ \psi(x)$$

We derive this as follows. Let us write $B\phi(x) = \chi(x)$. Then

$$\begin{aligned} \int_{-\infty}^{\infty} dx (AB\phi(x))^* \psi(x) &= \int_{-\infty}^{\infty} dx (A\chi(x))^* \psi(x) \\ &= \int_{-\infty}^{\infty} dx \chi^*(x) A^+ \psi(x) = \int_{-\infty}^{\infty} dx (B\phi(x))^* A^+ \psi(x) \\ &= \int_{-\infty}^{\infty} dx \phi^*(x) B^+ A^+ \psi(x) \end{aligned}$$

Since this is true for all $\psi(x)$ and $\phi(x)$, it follows that

$$(AB)^+ = B^+ A^+ \quad (5-31)$$

which is easily generalized to the rule

$$(ABC \dots Z)^+ = Z^+ \dots C^+ B^+ A^+ \quad (5-32)$$

4. For any operator Q , the combinations $Q + Q^+$, $i(Q - Q^+)$, and QQ^+ are hermitian.

5. The product of two hermitian operators is not necessarily hermitian, since

$$(AB)^+ = B^+ A^+ = BA = AB - (AB - BA)$$

The condition that it be hermitian is that the *commutator* of A and B , defined by

$$[A, B] = AB - BA$$

vanishes.

5-4 DEGENERACY AND SIMULTANEOUS OBSERVABLES

In both the problems discussed in Chapter 3, the particle in the box, and the free particle, we found that the eigenfunctions were simultaneous eigenfunctions of H and another operator—parity in the first case, momentum in the second—and we saw that in both cases the additional operators commuted with H . Let us now examine the general conditions under which this happens.

The eigenfunctions u_a , corresponding to the eigenvalue a of the operator A ,

$$Au_a(x) = au_a(x) \quad (5-33)$$

will be simultaneous eigenfunctions of another operator B , when

$$Bu_a(x) = bu_a(x) \quad (5-34)$$

This, however, implies that

$$ABu_a(x) = A bu_a(x) = bAu_a(x) = abu_a(x)$$

and

$$BAu_a(x) = B au_a(x) = aBu_a(x) = abu_a(x)$$

that is, that

$$(AB - BA)u_a(x) = 0 \quad (5-35)$$

If this were to hold for just one u_a , it would not be very interesting, but if it holds for the complete set u_a , then it means that for all square-integrable functions $\psi(x) = \sum_a C_a u_a(x)$,

$$\begin{aligned} \sum_a C_a (AB - BA) u_a(x) &= (AB - BA) \sum_a C_a u_a(x) \\ &= (AB - BA) \psi(x) = 0 \end{aligned} \quad (5-36)$$

that is, the operators commute

$$[A, B] = 0 \quad (5-37)$$

Does the converse apply? Suppose we have two hermitian operators that commute, so that (5-37) holds. Under these circumstances,

$$ABu_a(x) = BAu_a(x) = aBu_a(x)$$

Thus the function $Bu_a(x)$ is also an eigenfunction of A with eigenvalue a . If there is only one eigenfunction of A that corresponds to the eigenvalue a , then we must conclude that $Bu_a(x)$ is proportional to $u_a(x)$. We write this proportionality in the form

$$Bu_a(x) = bu_a(x) \quad (5-38)$$

and conclude that $u_a(x)$ is *simultaneously* an eigenfunction of both A and B . It should therefore more appropriately be labeled $u_{ab}(x)$, with

$$\begin{aligned} Au_{ab}(x) &= au_{ab}(x) \\ Bu_{ab}(x) &= bu_{ab}(x) \end{aligned} \quad (5-39)$$

We saw just this situation in the case of the infinite well, in which A was the Hamiltonian and B the parity operator, and in the case of the free particle, in which A was the Hamiltonian and B the momentum operator.

The above was obtained under the assumption that there was only one eigenfunction of A , with eigenvalue a . What happens when there is *degeneracy*; that is, there are several eigenfunctions of A with the same eigenvalue a ? In that case we can no longer conclude that (5-38) holds. Suppose that we have

$$\begin{aligned} Au_a^{(1)}(x) &= au_a^{(1)}(x) \\ Au_a^{(2)}(x) &= au_a^{(2)}(x) \end{aligned} \quad (5-40)$$

Under these circumstances we can only say that B acting on either of the functions $u_a^{(1)}(x)$, $u_a^{(2)}(x)$ will yield a linear combination of these. This reads as follows

$$\begin{aligned} Bu_a^{(1)} &= b_{11}u_a^{(1)} + b_{12}u_a^{(2)} \\ Bu_a^{(2)} &= b_{21}u_a^{(1)} + b_{22}u_a^{(2)} \end{aligned} \quad (5-41)$$

This presents no problem. Any linear combination of $u_a^{(1)}(x)$ and $u_a^{(2)}(x)$, which we will call $v_a^{(1)}(x)$, $v_a^{(2)}(x)$ will satisfy (5-40). We choose a linear combination such that

$$\begin{aligned} Bv_a^{(1)}(x) &= b_1v_a^{(1)}(x) \\ Bv_a^{(2)}(x) &= b_2v_a^{(2)}(x) \end{aligned} \quad (5-42)$$

(See Problem 14.) We again take into account that these are common eigenfunctions of A and B , and write this out in the form

$$\begin{aligned}Av_{ab} &= av_{ab} \\ Bv_{ab} &= bv_{ab}\end{aligned}\tag{5-43}$$

where b can take on two different values. Note that the eigenfunctions for these two values of b are orthogonal to each other, since they correspond to different eigenvalues b . In practice, for two-fold degeneracy such as we found in the two examples cited above, just choosing the two degenerate eigenfunctions of A that are *orthogonal to each other* (e.g., e^{ikx} and e^{-ikx} for the free particle case) will automatically be eigenfunctions of B .

It sometimes happens that even after this is done, there may still be several eigenfunctions that obey (5-40). This is a signal that there is yet a third operator C , which commutes with both A and B , and which may distinguish between these eigenfunctions. Under those circumstances we end up with equations like

$$\begin{aligned}Aw_{abc} &= aw_{abc} \\ Bw_{abc} &= bw_{abc} \\ Cw_{abc} &= cw_{abc}\end{aligned}\tag{5-44}$$

We have no example of such a situation in one-dimensional wave mechanics, but such a situation arises in three-dimensional problems—for example, in the case of the hydrogen atom discussed in Chapter 8. There one finds that there are three operators that commute with each other. These are H , the Hamiltonian, L^2 , the square of the angular momentum, and L_z , the component of the angular momentum along the z -axis, and the eigenfunctions of the hydrogen atom carry three labels, indicating the eigenvalues of the three operators.

The process outlined above goes on until there is no more degeneracy. The set of *mutually commuting* operators is called the *complete set of commuting observables*. The eigenvalues of this set, $a, b, c, \dots$, is the largest amount of information that one can have about a system *all at once*. The reason is that if we consider a new observable Q , say, that is not a function of the commuting operators ($A, B, C, \dots$) then a measurement of Q will not give a sharp value. For example, in the case of the free particle in one dimension, we have H, p_{op} , and parity P . The last two don't commute, since

$$\begin{aligned}Pp_{\text{op}}f(x) &= -i\hbar P \frac{df(x)}{dx} = +i\hbar \frac{df(-x)}{dx} \\ p_{\text{op}}Pf(x) &= p_{\text{op}}f(-x) = -i\hbar \frac{df(-x)}{dx}\end{aligned}$$

Thus if we choose simultaneous eigenfunctions of H and P —namely, $\cos kx$ and $\sin kx$ —then p_{op} acting on these will only give a linear combination of eigenfunctions of p_{op} .

This lack of sharpness is related to the uncertainty relations. In Supplement 5-A [www.wiley.com/college/gasiorowicz] we show the following: If we define the *uncertainty* ΔA in the measurement of A by its dispersion,

$$(\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2\tag{5-45}$$

and similarly for ΔB , then it follows that

$$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4} \langle i[A, B] \rangle^2\tag{5-46}$$

For the operators p_{op} and x , this yields

$$(\Delta p \Delta x)^2 \geq \frac{1}{4} \langle i(-i\hbar) \rangle^2$$

or

$$\Delta p \Delta x \geq \frac{1}{2} \hbar \quad (5-47)$$

5-5 TIME DEPENDENCE AND THE CLASSICAL LIMIT

Let us now turn to the important question of the classical limit of quantum theory. To do this we must first study the time development of expectation values of operators. In general, the expectation value of an operator changes with time. It may change with time because the operator has an explicit time dependence—for example, the operator $x + p/m$ —and it also changes with time because the expectation value is taken with respect to a wave function that itself changes with time. If we write

$$\langle A \rangle_t = \int \psi^*(x, t) A \psi(x, t) dx \quad (5-48)$$

then

$$\begin{aligned} \frac{d}{dt} \langle A \rangle_t &= \int \psi^*(x, t) \frac{\partial A}{\partial t} \psi(x, t) dx \\ &\quad + \int \frac{\partial \psi^*(x, t)}{\partial t} A \psi(x, t) dx \\ &\quad + \int \psi^*(x, t) A \frac{\partial \psi(x, t)}{\partial t} dx \\ &= \left\langle \frac{\partial A}{\partial t} \right\rangle_t + \int \left(\frac{1}{i\hbar} H \psi(x, t) \right)^* A \psi(x, t) dx \\ &\quad + \int \psi^*(x, t) A \left(\frac{1}{i\hbar} H \psi(x, t) \right) dx \\ &= \left\langle \frac{\partial A}{\partial t} \right\rangle_t + \frac{i}{\hbar} \int \psi^*(x, t) H A \psi(x, t) dx \\ &\quad - \frac{i}{\hbar} \int \psi^*(x, t) A H \psi(x, t) dx \end{aligned}$$

that is,

$$\frac{d}{dt} \langle A \rangle_t = \left\langle \frac{\partial A}{\partial t} \right\rangle_t + \frac{i}{\hbar} \langle [H, A] \rangle_t \quad (5-49)$$

In the derivation we made use of the fact that H is a hermitian operator. We observe that if A has no explicit time dependence, then the change of the expectation value *for any state* is

$$\frac{d}{dt} \langle A \rangle_t = \frac{i}{\hbar} \langle [H, A] \rangle_t \quad (5-50)$$

If the operator commutes with H , then its expectation value is always constant; that is, we may say that *the observable is a constant of the motion*. If the Hamiltonian is one of the complete set of commuting observables, then all the others are constants of the motion.

Let us consider successively $A = x$ and $A = p$. We first have

$$\begin{aligned}\frac{d}{dt}\langle x \rangle &= \frac{i}{\hbar} \langle [H, x] \rangle \\ &= \frac{i}{\hbar} \left\langle \left[\frac{p^2}{2m} + V(x), x \right] \right\rangle\end{aligned}$$

Now x commutes with any function of x ,

$$[V(x), x] = 0 \quad (5-51)$$

so that we only have to calculate

$$\begin{aligned}[p^2, x] &= p[p, x] + [p, x]p \\ &= \frac{2\hbar}{i} p\end{aligned} \quad (5-52)$$

Thus we obtain

$$\frac{d}{dt}\langle x \rangle = \left\langle \frac{p}{m} \right\rangle \quad (5-53)$$

Next we have

$$\begin{aligned}\frac{d}{dt}\langle p \rangle &= \frac{i}{\hbar} \left\langle \left[\frac{p^2}{2m} + V(x), p \right] \right\rangle \\ &= -\frac{i}{\hbar} \langle [p, V(x)] \rangle\end{aligned} \quad (5-54)$$

since p^2 and p evidently commute. To evaluate the last commutator, we note that

$$\begin{aligned}pV(x)\psi(x) - V(x)p\psi(x) &= \frac{\hbar}{i} \frac{d}{dx} [V(x)\psi(x)] - \frac{\hbar}{i} V(x) \frac{d}{dx} \psi(x) \\ &= \frac{\hbar}{i} \frac{dV(x)}{dx} \psi(x)\end{aligned} \quad (5-55)$$

so that

$$[p, V(x)] = \frac{\hbar}{i} \frac{dV(x)}{dx} \quad (5-56)$$

and thus

$$\frac{d}{dt}\langle p \rangle_t = -\left\langle \frac{dV(x)}{dx} \right\rangle_t \quad (5-57)$$

We may combine (5-53) and (5-57) to obtain

$$m \frac{d^2}{dt^2} \langle x \rangle_t = -\left\langle \frac{dV(x)}{dx} \right\rangle_t \quad (5-58)$$

which looks very much like the equation of motion of a classical point particle in a potential $V(x)$

$$m \frac{d^2 x_{cl}}{dt^2} = - \frac{dV(x_{cl})}{dx_{cl}} \quad (5-59)$$

The only thing that keeps us from making the identification

$$x_{cl} = \langle x \rangle \quad (5-60)$$

is that

$$\left\langle \frac{dV}{dx} \right\rangle \neq \frac{d}{d\langle x \rangle} V(\langle x \rangle) \quad (5-61)$$

Under the circumstances where the preceding inequality becomes an approximate equality, the motion is essentially classical, as was first noted by Ehrenfest. This requires that the potential be a slowly varying function of its argument. If we write

$$F(x) = - \frac{dV(x)}{dx} \quad (5-62)$$

then

$$F(x) = F(\langle x \rangle) + (x - \langle x \rangle) F'(\langle x \rangle) + \frac{(x - \langle x \rangle)^2}{2!} F''(\langle x \rangle) + \dots$$

If the uncertainty $(\Delta x)^2 = \langle (x - \langle x \rangle)^2 \rangle$ is small, and the higher terms in the expansion can be neglected, then we have

$$\begin{aligned} \langle F(x) \rangle &\cong F(\langle x \rangle) + \langle x - \langle x \rangle \rangle F'(\langle x \rangle) \\ &\cong F(\langle x \rangle) \end{aligned} \quad (5-63)$$

It is indeed true that even for electrons and other subatomic particles, (5-63) can be valid. For macroscopic fields (5-63) is a good approximation, and this allows us to describe electron or proton orbits in an accelerator by means of classical equations of motion.

PROBLEMS

1. Use (5-27) to prove that for a hermitian operator A ,

$$\int_{-\infty}^{\infty} dx \phi^*(x) A \psi(x) = \int_{-\infty}^{\infty} dx (A \phi(x))^* \psi(x)$$

(Hint: Let $\Psi = \phi + \lambda \psi$ in (5-27) and use the fact that λ is an arbitrary complex number.)

2. Prove that if A and B are hermitian, so is $(A + B)^n$.
3. Show that if A is hermitian, then $\langle A^2 \rangle$ is a positive number.
4. Prove that if H is a hermitian operator, then $e^{iH} = \sum_{n=0}^{\infty} \frac{(iH)^n}{n!}$ is a hermitian conjugate to e^{-iH} .
5. An operator U is said to be unitary, if $UU^\dagger = U^\dagger U = 1$. Prove that if H is hermitian, then e^{iH} is unitary.

6. Prove the *Schwartz inequality*

$$\left(\int_{-\infty}^{\infty} dx \phi^*(x) \phi(x) \right) \left(\int_{-\infty}^{\infty} dx \psi^*(x) \psi(x) \right) \geq \left| \int_{-\infty}^{\infty} dx \phi^*(x) \psi(x) \right|^2$$

(Hint: Use the fact that $\int_{-\infty}^{\infty} dx |\Psi(x)|^2 \geq 0$ with $\Psi(x) = \phi(x) + \lambda \psi(x)$ and find the minimum value. Note that the Schwartz inequality is the counterpart of the statement that $1 \geq \cos^2 \theta$ for the angle θ between any two three-dimensional vectors.)

7. Show that if two operators U and V are unitary, so is the operator UV .
8. Show that the eigenvalue λ of a unitary operator U must be of the form $e^{i\alpha}$. (Hint: Write

$$\int_{-\infty}^{\infty} dx (U\psi(x))^* U\psi(x) \text{ in two different ways.)}$$

9. Show that if $\psi(x)$ is a normalized wave function, and U is a unitary operator, then the function

$$\phi(x) = U\psi(x)$$

is also normalized to unity.

10. Consider a complete set of orthogonal, normalized eigenfunctions of some operator A denoted by $u_a(x)$. Given a unitary operator U we may construct the set $v_a(x)$ defined by $v_a(x) = Uu_a(x)$. Show that the new set of eigenfunctions is also orthonormal; that is,

$$\int_{-\infty}^{\infty} dx v_a^*(x) v_b(x) = \delta_{ab}$$

11. Operators that do not commute obey a number of relationships. Prove the following:
- (a) If A and B are hermitian, then $i[A, B]$ is also hermitian.
- (b) $[AB, C] = A[B, C] + [A, C]B$
- (c) $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$ (This is known as the Jacobi identity.)
12. Prove by expanding the exponentials that

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2!} [A, [A, B]] + \frac{1}{3!} [A, [A, [A, B]]] + \dots$$

13. Consider the hermitian operator H that has the property

$$H^4 = 1$$

What are the eigenvalues of the operator H ? What are the eigenvalues of H if it is not restricted to being hermitian?

14. Consider eq. (5-41). Find the linear combinations of $u_a^{(1)}$, $u_a^{(2)}$ that lead to the form (5-42), and find the eigenvalues b_1 and b_2 .
15. Use the commutation relations between the operators x and p to obtain the equations describing the time dependence of $\langle x \rangle$ and $\langle p \rangle$ for the Hamiltonian given by

$$H = \frac{p^2}{2m} + \frac{1}{2} m(\omega_1^2 x^2 + \omega_2 x + C)$$

16. Solve the equations of motion you obtained in Problem 15. Write your solutions in terms of $\langle x \rangle_0$ and $\langle p \rangle_0$, the expectation values at time $t = 0$.
17. An electron in an oscillating electric field is described by the Hamiltonian operator

$$H = \frac{p^2}{2m} - (eE_0 \cos \omega t)x$$

Calculate expressions for the time dependence of $\langle x \rangle$, $\langle p \rangle$, and $\langle H \rangle$.

18. Solve the equations of motion you obtained in Problem 17. Write your solutions in terms of $\langle x \rangle_0$ and $\langle p \rangle_0$, the expectation values at time $t = 0$.

Supplement 5-A

Uncertainty Relations

In our discussion of wave packets in Chapter 2, we noted that there is a relationship between the *spread* of a function and its Fourier transform. When the de Broglie correspondence between wave number and momentum is made, the relationship takes the form

$$\Delta p \Delta x \geq \hbar$$

What we called the *spread* or, in the above context, the uncertainty can be sharpened mathematically to a definition: The uncertainty in any physical variable ΔA is equal to the *dispersion*, given by

$$(\Delta A)^2 \equiv \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 \quad (5A-1)$$

With this definition we can show that the uncertainty relation follows directly from quantum mechanics. Let us consider two hermitian operators A and B , and shift them by a constant, their expectation values in an arbitrary physical state $\psi(x)$, so that we have

$$U = A - \langle A \rangle \quad (5A-2)$$

and

$$V = B - \langle B \rangle \quad (5A-3)$$

where

$$\langle A \rangle = \int_{-\infty}^{\infty} dx \psi^*(x) A \psi(x)$$

and so on. Given an arbitrary wave function $\psi(x)$, let us define

$$\phi(x) = (U + i\lambda V)\psi(x) \quad (5A-4)$$

with λ real. Whatever this function is, it will be true that

$$I(\lambda) = \int_{-\infty}^{\infty} dx \phi^*(x) \phi(x) \geq 0 \quad (5A-5)$$

This means, because of the hermiticity of U and V , that

$$\begin{aligned} I(\lambda) &= \int_{-\infty}^{\infty} dx ((U + i\lambda V)\psi(x))^* (U + i\lambda V)\psi(x) \\ &= \int_{-\infty}^{\infty} dx \psi^*(x) (U^+ - i\lambda V^+) (U + i\lambda V) \psi(x) \\ &= \int_{-\infty}^{\infty} dx \psi^*(x) (U - i\lambda V) (U + i\lambda V) \psi(x) \end{aligned}$$

It follows that

$$\begin{aligned} I(\lambda) &= \int_{-\infty}^{\infty} dx \psi^*(x)(U^2 + i\lambda[U, V] + \lambda^2 V^2)\psi(x) \\ &= \langle U^2 \rangle + i\lambda \langle [U, V] \rangle + \lambda^2 \langle V^2 \rangle \end{aligned} \quad (5A-6)$$

This will have its minimum value when

$$\frac{dI(\lambda)}{d\lambda} = 0 \quad (5A-7)$$

that is, when

$$i\langle [U, V] \rangle + 2\lambda \langle V^2 \rangle = 0$$

When

$$\lambda_{\min} = -\frac{i\langle [U, V] \rangle}{2\langle V^2 \rangle} \quad (5A-8)$$

is substituted into equation (5A-6) in the form

$$I(\lambda_{\min}) \geq 0 \quad (5A-9)$$

we get

$$\langle U^2 \rangle \langle V^2 \rangle \geq \frac{1}{4} \langle i[U, V] \rangle^2 \quad (5A-10)$$

or equivalently,

$$(\Delta A)^2 (\Delta B)^2 \geq \frac{1}{4} \langle i[A, B] \rangle^2 \quad (5A-11)$$

For the operators p and x for which

$$[p, x] = -i\hbar \quad (5A-12)$$

this leads to

$$\Delta p \Delta x \geq \frac{\hbar}{2} \quad (5A-13)$$

Note that if for $\psi(x)$ we take an eigenstate of the operator A , for example, then

$$\begin{aligned} (\Delta A)^2 &= \int_{-\infty}^{\infty} dx u_a^*(x) A^2 u_a(x) - \left(\int_{-\infty}^{\infty} dx u_a^*(x) A u_a(x) \right)^2 \\ &= a^2 \int_{-\infty}^{\infty} dx u_a^*(x) u_a(x) - \left(a \int_{-\infty}^{\infty} dx u_a^*(x) u_a(x) \right)^2 = 0 \end{aligned}$$

There is no problem, since the right side of the equation also vanishes:

$$\begin{aligned} \langle [A, B] \rangle &= \int_{-\infty}^{\infty} dx u_a^*(x) (AB - BA) u_a(x) = \int_{-\infty}^{\infty} dx (A u_a(x))^* B u_a(x) - B A u_a(x) \\ &= a \int_{-\infty}^{\infty} dx u_a^*(x) B u_a(x) - \int_{-\infty}^{\infty} dx u_a^*(x) B a u_a(x) = 0 \end{aligned}$$

We stress again that in this derivation no use was made of wave properties, x -space or p -space wave functions, or particle-wave duality. Our result depends entirely on the operator properties of the observables A and B .

Chapter 6

Operator Methods in Quantum Mechanics¹

In Chapter 5, in our discussion of the general principles of wave mechanics, we drew attention to the strong resemblance between the properties of wave functions and vectors in some sort of linear space, real or complex, finite- or infinite-dimensional. In this chapter we will restate this correspondence and make more abstract statements about quantum mechanics that go beyond wave mechanics and the Schrödinger equation. We shall work with a concrete example, the *simple harmonic oscillator*, and we shall also introduce a powerful notation, the *Dirac notation*, which provides a compact way of writing things and simplifies manipulations. Needless to say, when a number such as the wavelength of a photon emitted in a transition between states needs to be calculated, we must go back from the abstract notions to concrete manipulations. The advantage of the more abstract approach is that it allows us to deal with microscopic properties that do not have a large scale counterpart, and for which there is no *correspondence limit* to sharpen our intuition.

6-1 REPHRASING WAVE MECHANICS—AN ABSTRACT VIEW OF QUANTUM MECHANICS

We begin with the notion of a state in quantum mechanics. This may correspond to the n^{th} eigenstate of a particle in an infinite well, or it may describe a free particle traveling with momentum p from $x = -\infty$ toward a square barrier. All such states are described as *vectors* in some linear space. We shall describe the vectors by something called a *ket* and described by $|n\rangle$, or $|p\rangle$, or, for simultaneous eigenstates of two commuting operators, $|a, b\rangle$. In each case the label gives the value of the “observables” for the system in question. If the vector represents some other state that may be a superposition of eigenstates, then we denote this by $|\psi\rangle$. These *kets* have the properties of vectors. For example, the sum of two *kets*, with arbitrary complex numerical coefficients is again a *ket*. We will find a link between the abstract vector $|\psi\rangle$ and the wave function $\psi(x)$. Associated with the space of *kets* is a space of conjugate states, described by *bras*, and denoted by $\langle a|$. For each *ket* we have a *bra*. The correspondence between *kets* and wave functions will be matched by a similar correspondence between *bras* and complex conjugates of the wave functions. This will be explained in what follows.

¹This chapter deals with a generalization of the concepts introduced in Chapter 5. It can be read after Chapters 7 and 8, in which we still deal with the Schrödinger equation reduced, by the separation of variables, to a one-dimensional equation.

Consider next the *scalar product* of two eigenfunctions of some operator. We found that the eigenstates of the particle in the infinite well were orthogonal for different values of the energy; that is, we found

$$\int_{-\infty}^{\infty} dx u_m^*(x) u_n(x) = \delta_{mn} \quad (6-1)$$

How would we write the corresponding quantity in terms of *bras* and *kets*? We shall *define* the scalar product of a *bra* and a *ket* by the bracket (*bra-ket*) symbol $\langle m|n \rangle$. More generally, we define the scalar product of a *bra* $\langle \phi|$ and a *ket* $|\psi\rangle$ by the symbol $\langle \phi|\psi \rangle$. We will soon see that numerically it is identical to the familiar wave-mechanics expression

$$\langle \phi|\psi \rangle = \int_{-\infty}^{\infty} dx \phi^*(x) \psi(x) \quad (6-2)$$

We will also require that

$$\langle \phi|\psi \rangle^* = \langle \psi|\phi \rangle \quad (6-3)$$

and

$$\langle \phi|\alpha\psi_1 + \beta\psi_2 \rangle = \alpha\langle \phi|\psi_1 \rangle + \beta\langle \phi|\psi_2 \rangle \quad (6-4)$$

When an operator acts on a state it gives another state. In the Dirac language, this says that when an operator acts on a *ket* it gives another *ket*. We write this in two ways:

$$A|\psi \rangle = |A\psi \rangle \quad (6-5)$$

The left-hand side states this in the form of an operator acting on a *ket*, while the right-hand side is just our notation of the resulting *ket*. Thus

$$\langle \phi|A\psi \rangle = \langle \phi|A|\psi \rangle \quad (6-6)$$

Once we establish the equivalence (6-2) we will immediately get

$$\langle \phi|A|\psi \rangle = \int_{-\infty}^{\infty} dx \phi^*(x) A\psi(x) \quad (6-7)$$

In wave mechanics we defined the hermitian conjugate A^+ of the operator A by

$$\int_{-\infty}^{\infty} dx (A\phi(x))^* \psi(x) = \int_{-\infty}^{\infty} dx \phi(x)^* A^+ \psi(x) \quad (6-8)$$

This will have to emerge from the definition

$$\langle A\phi|\psi \rangle = \langle \phi|A^+|\psi \rangle \quad (6-9)$$

It follows that

$$\langle \phi|A^+|\psi \rangle^* = \langle A\phi|\psi \rangle^* = \langle \psi|A|\phi \rangle \quad (6-10)$$

Let us now consider the expansion theorem in the vector-space framework. In wave mechanics we saw that an arbitrary wave function could be expanded in terms of a complete set of eigenfunctions of a hermitian operator. An example of this was the expansion in terms of the eigenstates of the Hamiltonian for an infinite box. We mentioned an analogy

to the expansion of an arbitrary vector in terms of all the orthogonal unit vectors in the vector space. Here we start from the other end. We expand an arbitrary *ket* $|\psi\rangle$ in terms of some *complete set of orthonormal eigenstates*, which we label by $|n\rangle$:

$$|\psi\rangle = \sum_n C_n |n\rangle \quad (6-11)$$

It follows from the orthonormality condition $\langle m|n\rangle = \delta_{mn}$ that

$$C_n = \langle n|\psi\rangle \quad (6-12)$$

If we now substitute this into (6-11) we get

$$|\psi\rangle = \sum_n |n\rangle \langle n|\psi\rangle \quad (6-13)$$

where the sum is over the *complete set of eigenstates* of some hermitian operator. Since this must be true for every *ket* $|\psi\rangle$, it follows that

$$\sum_n |n\rangle \langle n| = \mathbf{1} \quad (6-14)$$

where $\mathbf{1}$ is the unit operator, which has the property that when it acts on any state, it does not change it. The relation (6-14) is called the *completeness relation*.

Before proceeding, we do an example in the manipulation of *bras* and *kets*.

EXAMPLE 6-1

Show that the eigenkets of any hermitian operator are orthogonal to each other if the eigenvalues are different.

SOLUTION The eigenvalue equations read

$$H|a\rangle = a|a\rangle$$

and

$$H|b\rangle = b|b\rangle$$

The hermiticity of H implies that the values of a and b are real. Let us now multiply the first equation by $\langle b|$. This gives us

$$\langle b|H|a\rangle = a\langle b|a\rangle$$

The second of the eigenvalue equations, multiplied by $\langle a|$, yields

$$\langle a|H|b\rangle = b\langle a|b\rangle$$

The complex conjugate of this reads

$$\langle a|H|b\rangle^* = \langle Hb|a\rangle = \langle b|H^+|a\rangle = b^*\langle a|b\rangle^* = b^*\langle b|a\rangle$$

However, because $H = H^+$ and b is real, this implies that

$$\langle b|H|a\rangle = b\langle b|a\rangle$$

It follows that

$$(a - b)\langle b|a\rangle = 0$$

A consequence of this is that if $a \neq b$ then $\langle b|a\rangle = 0$, which is what we set out to prove.

Let us return to the expansion theorem. In the expansion in (6-11) the interpretation of C_n is still that $|C_n|^2$ is the probability that a measurement of the observable A , whose eigenkets are the $|n\rangle$, yields the eigenvalue labeled by n . We wrote things out as if the set of $|n\rangle$ were discrete, but it need not be. The variables labeled by n could be continuous. Suppose we ask: What is the probability that if we make a measurement of the *position* of the state represented by $|\psi\rangle$, we find the value x ?

Since the position is an observable, we have every reason to expect that it is represented by a hermitian operator, which we denote by X . This operator will have a complete set of *eigenkets*, whose eigenvalues are the numbers x . We write this in the usual way:

$$X|x\rangle = x|x\rangle \quad (6-15)$$

We don't need to know how to construct the operator X . The expansion theorem now tells us that we can write

$$|\psi\rangle = \int_{-\infty}^{\infty} dx C(x)|x\rangle \quad (6-16)$$

with the sum replaced by an integral, since x is a continuously varying quantity. Using the orthonormality of the $|x\rangle$ eigenstates, in the form

$$\langle x|x'\rangle = \delta(x - x') \quad (6-17)$$

we calculate $C(x')$ by multiplying (6-16) from the left by $\langle x'|$,

$$C(x') = \langle x'|\psi\rangle \quad (6-18)$$

Now $|C(x)|^2$ is the probability that a measurement of the position of the state $|\psi\rangle$ yields the value x . This, however is just the property attributed to $|\psi(x)|^2$! All we have to do is change notation and write

$$\psi(x) = \langle x|\psi\rangle \quad (6-19)$$

A similar argument would show that the momentum space wave function for the state $|\psi\rangle$ is

$$\phi(p) = \langle p|\psi\rangle \quad (6-20)$$

We have symbolized this amplitude by the letter ϕ , because that is the notation we used earlier for the momentum space wave function.

We can now make several points:

(a) The completeness relation in terms of the position eigenkets reads

$$\int_{-\infty}^{\infty} dx |x\rangle\langle x| = \mathbf{1} \quad (6-21)$$

The unit operator can now be inserted at will. For example,

$$\langle\psi_1|\psi_2\rangle = \langle\psi_1|\mathbf{1}|\psi_2\rangle = \int_{-\infty}^{\infty} dx \langle\psi_1|x\rangle\langle x|\psi_2\rangle = \int_{-\infty}^{\infty} dx \psi_1^*(x)\psi_2(x) \quad (6-22)$$

which is what we asserted in (6-2). In the same way, (6-9) takes the form (6-8) because

$$\begin{aligned}\langle A\phi|\psi\rangle &= \int_{-\infty}^{\infty} dx (A\phi(x))^* \psi(x) \\ \langle \phi|A^+|\psi\rangle &= \int_{-\infty}^{\infty} \phi^*(x) A^+ \psi(x)\end{aligned}$$

(b) Consider

$$\psi(x) = \langle x|\psi\rangle = \int_{-\infty}^{\infty} dp \langle x|p\rangle \langle p|\psi\rangle = \int_{-\infty}^{\infty} dp \langle x|p\rangle \phi(p) \quad (6-23)$$

Comparison with

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \phi(p) e^{ipx/\hbar} \quad (6-24)$$

yields

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \quad (6-25)$$

Projection Operators

Let us return to our interpretation of the expansion coefficients. For simplicity we shall deal with a system in which all the eigenvalues are discrete, as for the infinite box. We then have

$$|\psi\rangle = \sum_n |n\rangle \langle n|\psi\rangle \quad (6-26)$$

The operator P_n defined by

$$P_n = |n\rangle \langle n| \quad (6-27)$$

has the properties that

$$P_m P_n = |m\rangle \langle m|n\rangle \langle n| = \delta_{mn} |m\rangle \langle n| = \delta_{mn} P_n \quad (6-28)$$

and

$$\sum_n P_n = 1 \quad (6-29)$$

In particular, it follows from (6-28) that

$$P_n^2 = P_n \quad (6-30)$$

We call P_n a *projection operator*, since it has the property that when acting on an arbitrary state $|\psi\rangle$ it *projects* it into the state $|n\rangle$, with a probability amplitude $\langle n|\psi\rangle$. Equation (6-28) states that once a state is projected into a particular eigenstate $|n\rangle$, then a further projection into that state does not change anything. This property of P_n fits in with our description of what a measurement is. We measure the energy of a state $|\psi\rangle$, for example. What this means is that of a collection of systems all of which are described by the same vector $|\psi\rangle$, we pick one and measure its energy. On finding a particular value E_n (with some prob-

ability), the measurement must change the system so that it ends up in the state $|n\rangle$. This is so because subsequent measurements in that state can only give the same result over and over again. A measurement of the energy of another member of this collection may yield a totally different eigenvalue— E_m , say—and in this way, the average energy becomes

$$\langle\psi|H|\psi\rangle = \sum_n |\langle n|\psi\rangle|^2 E_n = \sum_n \langle\psi|n\rangle E_n \langle n|\psi\rangle \quad (6-31)$$

This allows us to write the operator H in terms of its eigenvalues and the corresponding projectors

$$H = \sum_n E_n P_n \quad (6-32)$$

We will have further occasion to use the general description of states in our discussion of the harmonic oscillator below.

6-2 THE ENERGY SPECTRUM OF THE HARMONIC OSCILLATOR

The Hamiltonian for the harmonic oscillator has the form

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2 \quad (6-33)$$

where x and p are operators. The only properties of these operators that we shall use are that they are hermitian and that they obey the commutation relations

$$[p, x] = -i\hbar \quad (6-34)$$

Classically the Hamiltonian could be factored into

$$H = \omega \left(\sqrt{\frac{m\omega}{2}} x - i \frac{p}{\sqrt{2m\omega}} \right) \left(\sqrt{\frac{m\omega}{2}} x + i \frac{p}{\sqrt{2m\omega}} \right)$$

but because x and p do not commute, we have

$$\begin{aligned} \omega \left(\sqrt{\frac{m\omega}{2}} x - i \frac{p}{\sqrt{2m\omega}} \right) \left(\sqrt{\frac{m\omega}{2}} x + i \frac{p}{\sqrt{2m\omega}} \right) &= \frac{p^2}{2m} + \frac{m\omega^2}{2} x^2 - \frac{i\omega}{2} (px - xp) \\ &= H - \frac{1}{2} \hbar\omega \end{aligned} \quad (6-35)$$

We now introduce a special notation for the operators in terms of which $H - \frac{1}{2} \hbar\omega$ is factored. We write

$$\begin{aligned} A &= \sqrt{\frac{m\omega}{2\hbar}} x + i \frac{p}{\sqrt{2m\omega\hbar}} \\ A^+ &= \sqrt{\frac{m\omega}{2\hbar}} x - i \frac{p}{\sqrt{2m\omega\hbar}} \end{aligned} \quad (6-36)$$

with the additional factor $\sqrt{1/\hbar}$ introduced to make A and A^+ dimensionless. Since x and p are hermitian, A^+ is indeed the hermitian conjugate of A . We compute

$$[A, A^+] = \left[\sqrt{\frac{m\omega}{2\hbar}} x, -i \frac{p}{\sqrt{2m\omega\hbar}} \right] + \left[i \frac{p}{\sqrt{2m\omega\hbar}}, \sqrt{\frac{m\omega}{2\hbar}} x \right] = 1 \quad (6-37)$$

The Hamiltonian now takes the form

$$H = \hbar\omega \left(A^+ A + \frac{1}{2} \right) \quad (6-38)$$

The simplicity of the Hamiltonian is reflected in the simplicity of the commutation relations of A and A^+ with H . We have

$$\begin{aligned} [H, A] &= \hbar\omega [A^+ A, A] = -\hbar\omega A \\ [H, A^+] &= \hbar\omega [A^+ A, A^+] = \hbar\omega A^+ \end{aligned} \quad (6-39)$$

Let us now write down the eigenvalue equation, which reads

$$H|E\rangle = E|E\rangle \quad (6-40)$$

The first of the commutation relations in (6-39), acting on the state $|E\rangle$ yields

$$\begin{aligned} HA|E\rangle &= AH|E\rangle - \hbar\omega A|E\rangle \\ &= (E - \hbar\omega) A|E\rangle \end{aligned} \quad (6-41)$$

Thus $A|E\rangle$ is also an eigenstate of H , but with energy lowered by $\hbar\omega$. If we apply A again to the state $A|E\rangle$, we again lower the energy by $\hbar\omega$, so that we get the picture of a ladder going downward in energy. (See Fig. 6-1.) This lowering cannot go on forever, because the expectation value of H in any state is positive. To see this we observe that $\langle\psi|p^2|\psi\rangle = \langle p\psi|p\psi\rangle \geq 0$ and similarly $\langle\psi|x^2|\psi\rangle = \langle x\psi|x\psi\rangle \geq 0$, so that both terms in the Hamiltonian are larger than or equal to zero. This means that there is a state of lowest energy, the *ground state*. We denote it by $|0\rangle$. There is no implication that such a state has zero energy. We just use that label for the ground state, which must satisfy

$$A|0\rangle = 0 \quad (6-42)$$

so that the energy cannot be lowered any more. It follows from this and (6-38) that in fact the energy of the ground state is actually given by

$$H|0\rangle = \hbar\omega \left(A^+ A + \frac{1}{2} \right) |0\rangle = \frac{1}{2} \hbar\omega |0\rangle \quad (6-43)$$

The existence of a *zero-point energy* is forced upon us by the uncertainty relation. For any state for which $\langle p \rangle = 0$ and $\langle x \rangle = 0$, the expectation value of the Hamiltonian may be written in terms of $(\Delta p)^2 = \langle p^2 \rangle$ and $(\Delta x)^2 = \langle x^2 \rangle$ as

$$\langle H \rangle = \frac{(\Delta p)^2}{2m} + \frac{1}{2} m\omega^2 (\Delta x)^2$$

and the uncertainty relation forbids $\Delta p = \Delta x = 0$.

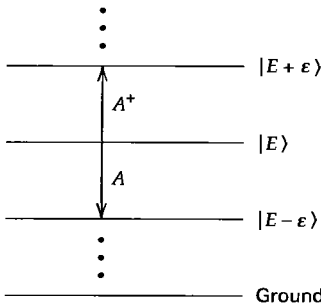


Figure 6-1 Raising and lowering operators, $\varepsilon = \hbar\omega$.

Consider now the state $A^+|0\rangle$. What is its energy? It follows from (6-39) that

$$HA^+|0\rangle = (A^+H + \hbar\omega A^+)|0\rangle = \hbar\omega\left(\frac{1}{2} + 1\right)A^+|0\rangle \quad (6-44)$$

This means that the state $A^+|0\rangle$ has one unit of energy $\hbar\omega$ more than the ground state. In the same way we show that a second application of A^+ to the state $A^+|0\rangle$ leads to a state of two units of energy above the ground state, and so on. From these conditions we conclude that the energy eigenvalues of the harmonic oscillator are

$$E = \left(n + \frac{1}{2}\right)\hbar\omega \quad (6-45)$$

The corresponding eigenket properly normalized is

$$|n\rangle = \frac{1}{\sqrt{n!}} (A^+)^n |0\rangle \quad (6-46)$$

To establish this we consider the product $AA^+A^+A^+ \dots A^+|0\rangle$ and use the commutation relation (6-37). We have

$$\begin{aligned} A(A^+)^n|0\rangle &= (A^+A + [A, A^+])(A^+)^{n-1}|0\rangle \\ &= (A^+)^{n-1}|0\rangle + A^+A(A^+)^{n-1}|0\rangle \\ &= (A^+)^{n-1}|0\rangle + A^+((A^+A + [A, A^+])(A^+)^{n-2}|0\rangle) \\ &= 2(A^+)^{n-1}|0\rangle + (A^+)^2((A^+A + [A, A^+])(A^+)^{n-3}|0\rangle) \\ &= 3(A^+)^{n-1}|0\rangle + (A^+)^3A(A^+)^{n-3}|0\rangle \\ &= \dots \\ &= n(A^+)^{n-1}|0\rangle + (A^+)^nA|0\rangle \\ &= n(A^+)^{n-1}|0\rangle \end{aligned} \quad (6-47)$$

We see that A acting on any polynomial $f(A^+)|0\rangle$ is equivalent to d/dA^+ acting on that state. Let us now consider $\langle 0|A^m(A^+)^n|0\rangle$. If $m > n$, then n of the A 's differentiate all the factors on the right, and we end up with $\langle 0|A^{m-n}|0\rangle = 0$. If $m < n$, then we end up with $\langle 0|(A^+)^{n-m}|0\rangle$ and this too vanishes. Finally, when $m = n$ we end up with

$$\langle 0|\left(\frac{d}{dA^+}\right)^n (A^+)^n|0\rangle = n(n-1)(n-2)\dots\langle 0|0\rangle = n! \quad (6-48)$$

which leads to (6-46) as well as

$$\langle m|n\rangle = \delta_{mn} \quad (6-49)$$

6-3 FROM OPERATORS BACK TO THE SCHRÖDINGER EQUATION

How do we get to the Schrödinger equation from our abstract framework? Consider the harmonic oscillator problem. Starting from (6-42), we get

$$\langle x|A|0\rangle = 0 \quad (6-50)$$

The original definition of A then implies that

$$\left\langle x \left| \sqrt{\frac{m\omega}{2\hbar}} x_{\text{op}} + i\sqrt{\frac{1}{2m\omega\hbar}} p_{\text{op}} \right| 0 \right\rangle = 0 \quad (6-51)$$

We now use

$$\langle x|x_{\text{op}}|0\rangle = x\langle x|0\rangle \quad (6-52)$$

and

$$\begin{aligned} \langle x|p_{\text{op}}|0\rangle &= \int_{-\infty}^{\infty} dp \langle x|p_{\text{op}}|p\rangle \langle p|0\rangle = \int_{-\infty}^{\infty} dp p \langle x|p\rangle \langle p|0\rangle \\ &= \int_{-\infty}^{\infty} dp p \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \langle p|0\rangle = \frac{\hbar}{i} \frac{d}{dx} \int_{-\infty}^{\infty} dp \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \langle p|0\rangle \quad (6-53) \\ &= \frac{\hbar}{i} \frac{d}{dx} \int_{-\infty}^{\infty} dp \langle x|p\rangle \langle p|0\rangle = \frac{\hbar}{i} \frac{d}{dx} \langle x|0\rangle \end{aligned}$$

We now denote $\langle x|0\rangle$ by $u_0(x)$, the ground-state wave function. Equation (6-50) now becomes a differential equation in x -space, which reads (after a little algebra)

$$\left(m\omega x + \hbar \frac{d}{dx}\right)u_0(x) = 0 \quad (6-54)$$

This is a simple differential equation, whose solution is

$$u_0(x) = Ce^{-m\omega x^2/2\hbar} \quad (6-55)$$

The constant C is determined by the normalization requirement that

$$1 = C^2 \int_{-\infty}^{\infty} dx e^{-m\omega x^2/\hbar} = C^2 \sqrt{\frac{\hbar\pi}{m\omega}} \quad (6-56)$$

so that

$$C = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \quad (6-57)$$

We may also obtain the higher energy states by working out in detail

$$\begin{aligned} u_n(x) &= \frac{1}{\sqrt{n!}} (A^+)^n u_0(x) \\ &= \frac{1}{\sqrt{n!}} \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \left(\sqrt{\frac{m\omega}{2\hbar}} x - \sqrt{\frac{\hbar}{2m\omega}} \frac{d}{dx}\right)^n e^{-m\omega x^2/2\hbar} \end{aligned} \quad (6-58)$$

In a general Schrödinger equation in which $H = \frac{p_{\text{op}}^2}{2m} + V(x_{\text{op}})$, we have

$$\langle x|V(x_{\text{op}})|E\rangle = V(x)\langle x|E\rangle \quad (6-59)$$

To get $\langle x|p_{\text{op}}|E\rangle$ we calculate

$$\begin{aligned} \langle x|p_{\text{op}}|E\rangle &= \int_{-\infty}^{\infty} dp \langle x|p_{\text{op}}|p\rangle \langle p|E\rangle = \int_{-\infty}^{\infty} dp p \langle x|p\rangle \langle p|E\rangle \\ &= \frac{\hbar}{i} \frac{d}{dx} \langle x|E\rangle \end{aligned} \quad (6-60)$$

by following the same steps that led to (6-53). From this we see that

$$\langle x|H|E\rangle = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \langle x|E\rangle + V(x)\langle x|E\rangle = E\langle x|E\rangle \quad (6-61)$$

which is just the Schrödinger energy eigenvalue equation!

6-4 THE TIME DEPENDENCE OF OPERATORS

We conclude this chapter by discussing the time development of a system in our representation-independent way. The equation that describes the evolution of a system is the time-dependent Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H|\psi(t)\rangle \quad (6-62)$$

This is now an operator equation in an abstract space. $|\psi(t)\rangle$ is a vector in that space and it points in a direction that depends on time. The solution of this equation is

$$|\psi(t)\rangle = e^{-iHt/\hbar} |\psi(0)\rangle \quad (6-63)$$

where $|\psi(0)\rangle$ is the vector at time $t = 0$. The operator in front is defined by

$$e^{-iHt/\hbar} = \sum_{n=0}^{\infty} \frac{1}{n!} (-iHt/\hbar)^n \quad (6-64)$$

The solution (6-63) allows us to describe the change with time of the expectation value of any operator B that does not have any explicit time dependence:

$$\begin{aligned} \langle B \rangle_t &= \langle \psi(t) | B | \psi(t) \rangle \\ &= \langle e^{-iHt/\hbar} \psi(0) | B | e^{-iHt/\hbar} \psi(0) \rangle \\ &= \langle \psi(0) | e^{iHt/\hbar} B e^{-iHt/\hbar} | \psi(0) \rangle \\ &= \langle \psi(0) | B(t) | \psi(0) \rangle \end{aligned} \quad (6-65)$$

where we have defined

$$B(t) = e^{iHt/\hbar} B e^{-iHt/\hbar} \quad (6-66)$$

What (6-65) states is that the expectation value of a time-independent operator B in a state that varies with time as given by (6-63) may be written as the expectation value of a time-varying operator $B(t)$ (given by (6-65)) in the time-independent state $|\psi(0)\rangle$. This is very useful in the formal discussion of quantum mechanics, since it is convenient to set up a basis of orthonormal eigenvectors in the abstract space once and for all, and not worry about how the basis vectors change with time. When we do this, we are working in the *Heisenberg picture*, whereas keeping B without time dependence means that we are working in the *Schrödinger picture*. The result is the same, whatever picture we use: This is analogous to the option of describing a rotating body relative to a fixed set of axes, or of describing the body at rest in a rotating coordinate system. The choice is one of convenience. If we do work in the Heisenberg picture, the state vectors are fixed, and we need not refer to them. How an observable varies with time is determined by (6-66), which yields

$$\begin{aligned} \frac{d}{dt} B(t) &= \frac{i}{\hbar} H e^{iHt/\hbar} B e^{-iHt/\hbar} - \frac{i}{\hbar} e^{iHt/\hbar} B e^{-iHt/\hbar} H \\ &= \frac{i}{\hbar} (HB(t) - B(t)H) = \frac{i}{\hbar} [H, B(t)] \end{aligned} \quad (6-67)$$

For the harmonic oscillator

$$H = \hbar\omega(A^+A + \frac{1}{2}) \quad (6-68)$$

and since H is a constant of the motion, we have

$$H = \hbar\omega(A^+(t)A(t) + \frac{1}{2}) \quad (6-69)$$

We can also show, using (6-66), that

$$[A(t), A^+(t)] = 1 \quad (6-70)$$

Eq. (6-67) implies that

$$\begin{aligned} \frac{d}{dt}A(t) &= -i\omega A(t) \\ \frac{d}{dt}A^+(t) &= i\omega A^+(t) \end{aligned} \quad (6-71)$$

whose solutions are

$$\begin{aligned} A(t) &= e^{-i\omega t}A(0) \\ A^+(t) &= e^{i\omega t}A^+(0) \end{aligned} \quad (6-72)$$

We may use these to express the operators $p(t)$ and $x(t)$ in terms of $p(0)$ and $x(0)$:

$$\begin{aligned} p(t) &= p(0) \cos \omega t - m\omega x(0) \sin \omega t \\ x(t) &= x(0) \cos \omega t + \frac{p(0)}{m\omega} \sin \omega t \end{aligned} \quad (6-73)$$

Note that in the Heisenberg picture we only deal with operators. We have therefore omitted the subscript op to avoid cluttering up the equations.

PROBLEMS

- (a) Show that the eigenvalues of a hermitian operator A (for which $A = A^+$) are real.
(b) Prove that $(AB)^+ = B^+A^+$.
(Hint: Use eqs. (6-5)–(6-10) to prove that $\langle\psi|(AB)^+|\psi\rangle = \langle\psi|B^+A^+|\psi\rangle$.)
- We may define the *trace* of an operator as follows:

$$\text{Tr } A = \sum_n \langle n|A|n\rangle$$

Where the sum is over a complete set of states. Prove that $\text{Tr } AB = \text{Tr } BA$. (Hint: Use the representation of $\mathbf{1}$ in terms of P_n .)

- Consider the harmonic oscillator. Prove that

$$A|n\rangle = \sqrt{n}|n-1\rangle$$

- Prove that if $f(A^+)$ is any polynomial in A^+ , then

$$Af(A^+)|0\rangle = \frac{df(A^+)}{dA^+}|0\rangle$$

[Hint: See Eq. (6-47).]

- Calculate $\langle m|x|n\rangle$ and show that it vanishes unless $n = m \pm 1$.
- Calculate $\langle m|p|n\rangle$.

7. Use the results of Problems 6 and 7 to calculate $\langle m|px|n\rangle$ and $\langle m|xp|n\rangle$. (*Hint:* Insert the unit operator **1** between the two operators, and use the form

$$\mathbf{1} = \sum_k |k\rangle\langle k|$$

In Problems 5–7 it is useful to express x and p in terms of A and $A^\dagger$

8. Use the results of Problem 7 to calculate $\langle m|[p, x]|n\rangle$.
9. Prove that $\langle n|x|n\rangle$ and $\langle n|p|n\rangle$ both vanish.
10. Calculate $\langle n|x^2|n\rangle$ and $\langle n|p^2|n\rangle$.
11. Use the definition of $(\Delta x)^2 = \langle n|x^2|n\rangle - \langle n|x|n\rangle^2$ and similarly for $(\Delta p)^2$ to calculate $\Delta x \Delta p$ for the state $|n\rangle$ of a harmonic oscillator.
12. A state $|\alpha\rangle$ that obeys the equation $A|\alpha\rangle = \alpha|\alpha\rangle$, where A is the harmonic oscillator annihilation operator, is called a *coherent state*.
- (a) Show that the state $|\alpha\rangle$ may be written in the form

$$|\alpha\rangle = C e^{\alpha A^\dagger} |0\rangle$$

- (b) Use the result of Problem 4 to obtain C .
- (c) Expand the state $|\alpha\rangle$ in a series of eigenstates of the operator $A^\dagger A$, $|n\rangle$, and use this to find the probability that the coherent state contains n quanta. The distribution is called a *Poisson distribution*.
- (d) Calculate $\langle \alpha|N|\alpha\rangle$, the average number of quanta in the coherent state, where $N = A^\dagger A$.
13. Use the general operator equation of motion (6-67) to solve for the time dependence of the operator $x(t)$ given that

$$H = \frac{p^2(t)}{2m} + mgx(t)$$

14. Consider the Hamiltonian describing a one-dimensional oscillator in an external electric field

$$H = \frac{p^2(t)}{2m} + \frac{1}{2}m\omega^2 x^2(t) - e\mathcal{E}x(t)$$

Calculate the equation of motion for the operators $p(t)$ and $x(t)$ using (6-67) and the commutation relation

$$[p(t), x(t)] = \frac{\hbar}{i}$$

Show that the equation of motion is just the classical equation of motion. Solve for $p(t)$ and $x(t)$ in terms of $p(0)$ and $x(0)$. Show that

$$[x(t_1), x(t_2)] \neq 0 \quad \text{for } t_1 \neq t_2$$

This shows that operators that commute at the same time need not commute at different times.

15. Use (6-58) to calculate the eigenfunctions for $n = 1, 2, 3$. (*Note:* Be sure to keep track of the ordering of x and d/dx in the expansion of the binomial series.)
16. Use the results of Problem 4 to show that

$$e^{\lambda A} f(A^\dagger) |0\rangle = f(A^\dagger + \lambda) |0\rangle$$

(*Hint:* Expand the exponential in a series, and use the fact that

$$f(x+a) = \sum \frac{a^n}{n!} f^{(n)}(x)$$

$$f^{(n)}(x) = \frac{d^n}{dx^n} f(x)$$

to work out this problem.)

17. Use the results of Problem 16 to establish the operator relation

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

Note that an operator relation must hold when it acts on an arbitrary state. Let an arbitrary state be of the form $g(A^\dagger)|0\rangle$. Thus what must be proved is that

$$e^{\lambda A} f(A^\dagger) e^{-\lambda A} g(A^\dagger)|0\rangle = f(A^\dagger + \lambda) g(A^\dagger)|0\rangle$$

This can also be proved from the general relation

$$e^{\lambda A} A^\dagger e^{-\lambda A} = A^\dagger + \lambda[A, A^\dagger] + \frac{\lambda^2}{2!}[A, [A, A^\dagger]] + \dots$$

18. Use the preceding relation to prove that

$$e^{aA+bA^\dagger} = e^{aA} e^{bA^\dagger} e^{-(1/2)ab}$$

The procedure is the following. Let

$$e^{\lambda(aA+bA^\dagger)} \equiv e^{\lambda aA} F(\lambda)$$

Differentiation with respect to λ yields

$$(aA + bA^\dagger) e^{\lambda(aA+bA^\dagger)} = aA e^{\lambda aA} F(\lambda) + e^{\lambda aA} \frac{dF}{d\lambda}$$

that is,

$$(aA + bA^\dagger) e^{\lambda aA} F(\lambda) = aA e^{\lambda aA} (F\lambda) + e^{\lambda aA} \frac{dF}{d\lambda}$$

Use Problem 17 to show that

$$\frac{dF}{d\lambda} = (bA^\dagger - \lambda ab) F(\lambda)$$

so that

$$F(\lambda) = e^{\lambda bA^\dagger} e^{-(1/2)\lambda^2 ab}$$

19. Use the procedure of Problem 17 to show that

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

Show from this that

$$e^{aA+bA^\dagger} = e^{bA^\dagger} e^{aA} e^{(1/2)ab}$$

using the method outlined in Problem 18.

20. Use the preceding result to show that

$$e^{ikx} = e^{ik\sqrt{\hbar/2m\omega} A^\dagger} e^{ik\sqrt{\hbar/2m\omega} A} e^{-(\hbar k^2/4m\omega)}$$

Note that

$$x = \sqrt{\frac{\hbar}{2m\omega}} (A + A^\dagger)$$

Use this expression to calculate

$$\langle 0|e^{ikx}|0\rangle$$

21. Show that the result just obtained is the same as the one obtained from

$$\int_{-\infty}^{\infty} dx u_0^*(x) e^{ikx} u_0(x)$$

Chapter 7

Angular Momentum

In our discussion of the Schrödinger equation in three dimensions we will need to deal with the kinetic energy in three dimensions, which has the form

$$K = \frac{\mathbf{p}^2}{2\mu} \quad (7-1)$$

for problems in which the effective mass of the particle is μ . Just as in classical mechanics, there is a close connection between $\mathbf{p}^2$ and the square of the angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. We shall take over this expression for the angular momentum into quantum mechanics, with the recognition that $\mathbf{r}$ and $\mathbf{p}$ are to be treated as operators. We will also see (Supplement 7-A [www.wiley.com/college/gasiorowicz]) that when the potential in a three-dimensional Schrödinger equation

$$H\psi(\mathbf{r}) = \left(\frac{\mathbf{p}^2}{2\mu} + V(\mathbf{r}) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (7-2)$$

is central, so that

$$V(\mathbf{r}) = V(r) \quad (7-3)$$

then angular momentum is conserved.

The conservation of angular momentum implies the operator equation

$$\frac{d\mathbf{L}}{dt} = 0 \quad (7-4)$$

This, as seen in Chapter 6, is equivalent to

$$[H, \mathbf{L}] = 0 \quad (7-5)$$

Our discussion of the details of the eigenvalues and eigenfunctions of the angular momentum operators is presented in two ways. We may proceed from $\mathbf{p}^2$ and obtain the angular momentum eigenvalue equation directly as a differential equation, which can be solved in a standard way. The other way proceeds from the definition of the angular momentum operator, and deals with the angular momentum commutation relations. This approach is patterned after the way we dealt with the harmonic oscillator in Chapter 6, and is more directly “quantum mechanical.” We therefore relegate the direct approach to Supplement 7-B [www.wiley.com/college/gasiorowicz], and turn directly to the algebraic, operator approach.

7-1 THE ANGULAR MOMENTUM COMMUTATION RELATIONS

We might be tempted to look for simultaneous eigenfunctions of H and all three components of $\mathbf{L}$. This, as we have seen in Chapter 5, is only possible if all four operators commute with each other. To proceed we must check whether *all* of the operators L_x , L_y , and L_z commute with each other, as well as with H . In fact, different components of the angular momentum *do not* commute with each other. For example, paying particular attention to the ordering of the operators, we get, using $\mathbf{L} = \mathbf{r} \times \mathbf{p}$,

$$\begin{aligned} [L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= y[p_z, z]p_x + x[z, p_z]p_y \\ &= \frac{\hbar}{i} (yp_x - xp_y) = i\hbar L_z \end{aligned} \quad (7-6a)$$

Similarly we can show that

$$[L_y, L_z] = i\hbar L_x \quad (7-6b)$$

and

$$[L_z, L_x] = i\hbar L_y \quad (7-6c)$$

It is true that each of the components of angular momentum commutes with $\mathbf{L}^2$. For example,

$$\begin{aligned} [L_z, L_x^2 + L_y^2 + L_z^2] &= L_y[L_z, L_x] + [L_z, L_y]L_x + L_x[L_z, L_x] + [L_z, L_x]L_x \\ &= -i\hbar L_y L_x - i\hbar L_x L_y + i\hbar L_x L_y + i\hbar L_y L_x = 0 \end{aligned} \quad (7-7)$$

We can see that as a consequence of these commutation relations, only one component of $\mathbf{L}$ may be chosen with H and $\mathbf{L}^2$ to form a simultaneously commuting set. To show this let us assume that we have a set of eigenfunctions that are simultaneous eigenfunctions of all three components of $\mathbf{L}$. Let us assume that

$$L_x|u\rangle = l_1|u\rangle$$

and

$$L_y|u\rangle = l_2|u\rangle$$

which implies that $L_x L_y|u\rangle = l_1 l_2|u\rangle$ and $L_y L_x|u\rangle = l_2 l_1|u\rangle$. As a consequence of (7-6a) this means that $L_z|u\rangle = 0$. This, however implies that

$$l_2|u\rangle = L_y|u\rangle = \frac{1}{i\hbar} [L_z, L_x]|u\rangle = \frac{1}{i\hbar} L_z l_1|u\rangle = 0$$

Similarly, we can show that $l_1|u\rangle = 0$. This means that only for $\mathbf{L} = 0$ can we have simultaneous eigenfunction for all three components of the angular momentum.

There is nothing to keep us from picking just one component of $\mathbf{L}$ as part of the commuting set. Conventionally the choice is L_z , but there is nothing special about this choice. We thus will deal with simultaneous eigenfunctions of $\mathbf{L}^2$ and L_z . We will denote the eigenkets by $|l, m\rangle$. Our starting point is thus the set of equations

$$\begin{aligned} \mathbf{L}^2|l, m\rangle &= \hbar^2 l(l+1)|l, m\rangle \\ L_z|l, m\rangle &= \hbar m|l, m\rangle \end{aligned} \quad (7-8)$$

At this point we make no statements about what the values of l and m are. They must be real since the operators $\mathbf{L}$ are hermitian. The factors of $\hbar$ are there purely for dimensional reasons, since the product $\mathbf{r} \times \mathbf{p}$ has the dimensions of $\hbar$.

7-2 RAISING AND LOWERING OPERATORS FOR ANGULAR MOMENTUM

Our starting point is (7-8), together with the angular momentum commutation relations (7-6) and the orthonormality relation

$$\langle l', m' | l, m \rangle = \delta_{ll'} \delta_{mm'} \quad (7-9)$$

It will prove convenient to introduce the operators

$$L_{\pm} = L_x \pm iL_y \quad (7-10)$$

These obey the commutation relations

$$\begin{aligned} [L_+, L_-] &= [L_x + iL_y, L_x - iL_y] = (-2i)[L_x, L_y] \\ &= 2\hbar L_z \end{aligned} \quad (7-11)$$

and

$$\begin{aligned} [L_z, L_{\pm}] &= [L_z, L_x \pm iL_y] = i\hbar L_y \mp i(\hbar L_x) = \pm\hbar(L_x \pm iL_y) \\ &= \pm\hbar L_{\pm} \end{aligned} \quad (7-12)$$

It is also obvious that

$$[L^2, L_{\pm}] = 0 \quad (7-13)$$

Furthermore, we have

$$\begin{aligned} L_+ L_- &= (L_x + iL_y)(L_x - iL_y) = L_x^2 + L_y^2 - i[L_x, L_y] \\ &= L^2 - L_z^2 + \hbar L_z \end{aligned} \quad (7-14)$$

and similarly,

$$L_- L_+ = L^2 - L_z^2 - \hbar L_z \quad (7-15)$$

Thus

$$L_+ L_- + L_z^2 - \hbar L_z = L_- L_+ + L_z^2 + \hbar L_z = L^2 \quad (7-16)$$

We now note that $\langle l, m | L_x^2 | l, m \rangle = \langle L_x(l, m) | L_x(l, m) \rangle \geq 0$ and by extension $\langle l, m | L^2 | l, m \rangle \geq 0$. This implies that $l(l+1) \geq 0$. From this it follows that $l \geq 0$. (The alternative that $l \leq -1$ we reject, since we would then call $l+1 = -l'$ and get $l' \geq 0$.) First, we note that

$$L^2 L_{\pm} |l, m\rangle = L_{\pm} L^2 |l, m\rangle = \hbar^2 l(l+1) L_{\pm} |l, m\rangle \quad (7-17)$$

This means that $L_{\pm} |l, m\rangle$ is an eigenstate of L^2 with the eigenvalue characterized by l . On the other hand,

$$L_z L_+ |l, m\rangle = (L_+ L_z + \hbar L_+) |l, m\rangle = \hbar(m+1) L_+ |l, m\rangle \quad (7-18)$$

and similarly

$$L_z L_- |l, m\rangle = \hbar(m-1) L_- |l, m\rangle \quad (7-19)$$

These equations imply that $L_+ |l, m\rangle$ is an eigenstate of L_z with the m value raised by unity, and $L_- |l, m\rangle$ is an eigenstate of L_z with m value lowered by unity. We therefore call $L_{\pm}$ *raising* and *lowering* operators, respectively. We may write

$$\begin{aligned} L_+ |l, m\rangle &= C_+(l, m) |l, m+1\rangle \\ L_- |l, m\rangle &= C_-(l, m) |l, m-1\rangle \end{aligned} \quad (7-20)$$

The conjugate relation of the first of the above is

$$\langle l, m | L_- = \langle l, m + 1 | C_+^*(l, m) \quad (7-21)$$

Multiplying this with the first of (7-20) yields

$$\begin{aligned} |C_+(l, m)|^2 \langle l, m + 1 | l, m + 1 \rangle &= \langle l, m | L_- L_+ | l, m \rangle = \langle l, m | \mathbf{L}^2 - L_z^2 - \hbar L_z | l, m \rangle \\ &= \hbar^2 [l(l + 1) - m^2 - m] \\ &= \hbar^2 [(l - m)(l + m + 1)] \end{aligned} \quad (7-22)$$

Thus

$$C_+(l, m) = \hbar \sqrt{(l - m)(l + m + 1)} \quad (7-23)$$

and, similarly

$$C_-(l, m) = \hbar \sqrt{l(l + 1) - m(m - 1)} = \hbar \sqrt{(l + m)(l - m + 1)} \quad (7-24)$$

It follows from

$$\langle L_\pm(l, m) | L_\pm(l, m) \rangle \geq 0$$

that

$$\begin{aligned} \langle L_\pm(l, m) | L_\pm(l, m) \rangle &= \langle l, m | L_\mp L_\pm | l, m \rangle \\ &= \langle l, m | \mathbf{L}^2 - L_z^2 \pm \hbar L_z | l, m \rangle \\ &= \hbar^2 [l(l + 1) - m(m \mp 1)] \geq 0 \end{aligned} \quad (7-25)$$

This implies that both

$$\begin{aligned} l(l + 1) &\geq m(m + 1) \\ l(l + 1) &\geq m(m - 1) \end{aligned} \quad (7-26)$$

are true. Since $l \geq 0$, it follows from the above that

$$-l \leq m \leq l \quad (7-27)$$

Let us assume that the minimum value of m is $m_{\min}$. This means that we cannot lower the m -value any further, and thus

$$L_- |l, m_{\min}\rangle = 0 \quad (7-28)$$

We can see, in a number of ways (by looking at $C_-(l, m)$ for example), that

$$m_{\min} = -l \quad (7-29)$$

Similarly, the maximum value of m , denoted by $m_{\max}$ is such that

$$L_+ |l, m_{\max}\rangle = 0 \quad (7-30)$$

and

$$m_{\max} = l \quad (7-31)$$

Since the maximum value is to be reached from the minimum value by unit steps (repeated application of L_+), we find, as seen in Fig. 7-1, that there are $(2l + 1)$ steps. This implies that $(2l + 1)$ is an integer, and m can take on the values

$$m = -l, -l + 1, -l + 2, \dots, l - 1, l \quad (7-32)$$

The possibility that l is half-integral, $l = 1/2, 3/2, 5/2, \dots$, will be discussed in Chapter 10, when we discuss *spin*. Until then, we restrict ourselves to *integer values of l* .

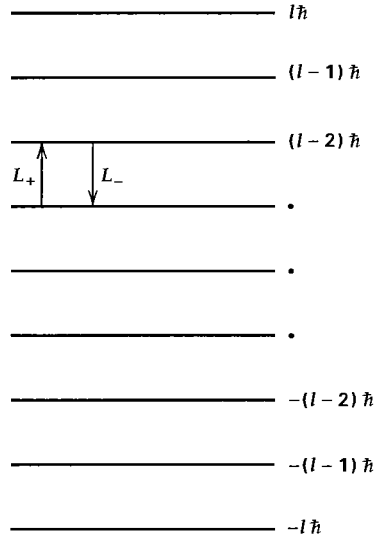


Figure 7-1 Spectrum of the operator L_z for a given value of l .

7-3 REPRESENTATION OF $|l, m\rangle$ STATES IN SPHERICAL COORDINATES

Our next task is to work out the general form of $Y_{lm}(\theta, \varphi)$, defined by

$$Y_{lm}(\theta, \varphi) = \langle \theta, \varphi | l, m \rangle \quad (7-33)$$

The angles (θ, φ) are the coordinates on the surface of a unit sphere (Fig. 7-2). The analog of

$$\int_{-\infty}^{\infty} dx |x\rangle \langle x| = \mathbf{1}$$

is

$$\int d\Omega |\theta, \varphi\rangle \langle \theta, \varphi| \equiv \int_0^{2\pi} d\varphi \int_0^\pi \sin \theta d\theta |\theta, \varphi\rangle \langle \theta, \varphi| = \mathbf{1} \quad (7-34)$$

It follows that the application of $\langle \theta', \varphi' |$ to (7-34) leads to

$$\langle \theta', \varphi' | \theta, \varphi \rangle = \frac{1}{\sin \theta} \delta(\theta - \theta') \delta(\varphi - \varphi') \quad (7-35)$$

To find an explicit form of the spherical harmonics as defined in (7-33) we make use of the polar coordinate representations

$$\begin{aligned} L_z &= \frac{\hbar}{i} \frac{\partial}{\partial \varphi} \\ L_{\pm} &= \hbar e^{\pm i\varphi} \left(\pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right) \end{aligned} \quad (7-36)$$

which are easily obtained from the material in Supplement 7-B.

Consider now $\langle \theta, \varphi | L_z | l, m \rangle$. This can be written in two ways. In one of them we have

$$\langle \theta, \varphi | L_z | l, m \rangle = \hbar m \langle \theta, \varphi | l, m \rangle \quad (7-37)$$

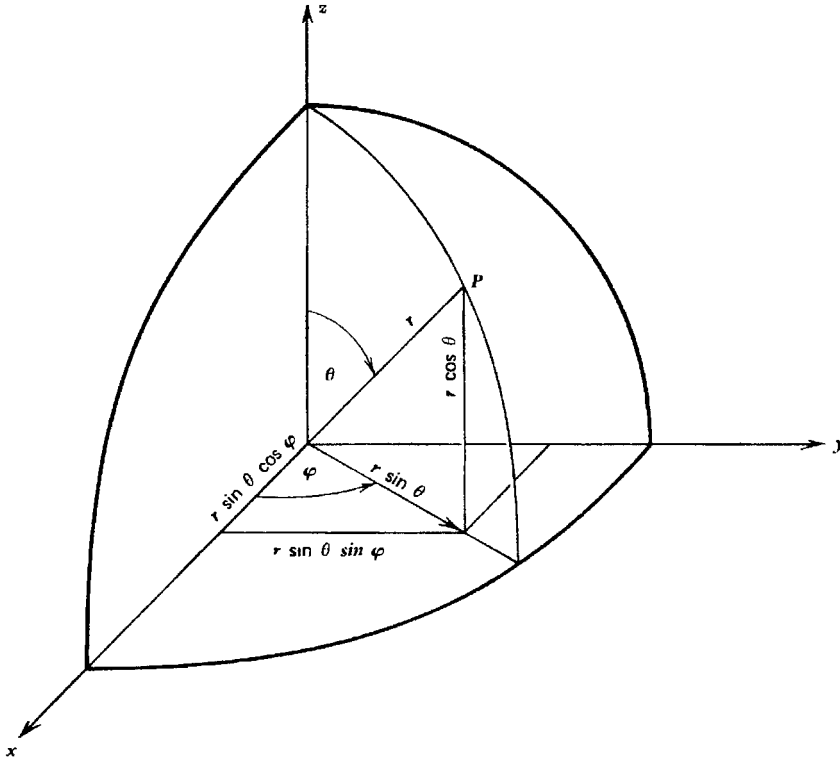


Figure 7-2 The definition of the spherical coordinates used in the text and the relation between the cartesian coordinates (x, y, z) and the spherical coordinates (r, θ, φ) .

In the other we use (7-36) to write

$$\begin{aligned} \frac{i}{\hbar} \langle \theta, \varphi | L_z | l, m \rangle &= \lim_{\delta\varphi \rightarrow 0} [\langle \theta, \varphi + \delta\varphi | l, m \rangle - \langle \theta, \varphi | l, m \rangle] / \delta\varphi \\ &= \frac{\partial}{\partial \varphi} \langle \theta, \varphi | l, m \rangle \end{aligned} \quad (7-38)$$

Combining (7-37) and (7-38) we get the differential equation

$$\frac{\partial}{\partial \varphi} \langle \theta, \varphi | l, m \rangle = im \langle \theta, \varphi | l, m \rangle \quad (7-39)$$

whose solution is

$$\langle \theta, \varphi | l, m \rangle = F(\theta) e^{im\varphi} \quad (7-40)$$

We next consider the maximum m -state $|l, l\rangle$ for which $L_+ |l, l\rangle = 0$. This implies that

$$\begin{aligned} 0 &= \langle \theta, \varphi | L_+ | l, l \rangle = \hbar e^{i\varphi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right) \langle \theta, \varphi | l, l \rangle \\ &= \hbar e^{i\varphi} \left(\frac{\partial}{\partial \theta} - l \cot \theta \right) \langle \theta, \varphi | l, l \rangle = \hbar e^{i(l+1)\varphi} \left(\frac{\partial}{\partial \theta} - l \cot \theta \right) F(\theta) \end{aligned} \quad (7-41)$$

or

$$\left(\frac{\partial}{\partial \theta} - l \cot \theta\right)F(\theta) = 0$$

The solution of this equation is easily seen to be

$$F(\theta) = (\sin \theta)^l \quad (7-42)$$

States with values of m lower than l can be obtained by repeated application of the *lowering operator* L_- to the state $|l, l\rangle$.¹ This amounts to

$$\begin{aligned} Y_{lm}(\theta, \varphi) &= C(L_-)^{l-m} (\sin \theta)^l e^{il\varphi} \\ &= C \left(-\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right)^{l-m} [(\sin \theta)^l e^{il\varphi}] \end{aligned} \quad (7-43)$$

The coefficient C is there so that the eigenstates are normalized according to

$$\int d\Omega |Y_{lm}(\theta, \varphi)|^2 = 1 \quad (7-44)$$

The details of working out the Y_{lm} can be found in Supplement 7-B. [www.wiley.com/college/gasiorowicz] Here we restrict ourselves to writing down the appropriately normalized Y_{lm} with the phases that are conventionally used. We have

$$Y_{lm}(\theta, \varphi) = (-1)^m \left[\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{1/2} P_l^m(\cos \theta) e^{im\varphi} \quad (7-45)$$

for $m \geq 0$, with

$$Y_{l,-m} = (-1)^m Y_{lm}^* \quad (7-46)$$

The *associated Legendre polynomials* are given by

$$P_l^m(u) = (-1)^{l+m} \frac{(l+m)!}{(l-m)!} \frac{(1-u^2)^{-m/2}}{2^l l!} \left(\frac{d}{du} \right)^{l-m} (1-u^2)^l \quad (7-47)$$

for $m \geq 0$. The values for negative m are obtained from

$$P_l^{-m}(u) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(u) \quad (7-48)$$

We list some of the spherical harmonics in the following table.

¹We may relate our results here to our discussion of *degeneracy* in Chapter 5. There are $(2l+1)$ eigenfunctions of $\mathbf{L}^2$. In other words, the eigenfunctions are $(2l+1)$ -fold degenerate. We found another operator L_z that commutes with $\mathbf{L}^2$, and we were able to arrange things so that the eigenfunctions are *simultaneous* eigenfunctions of the *mutually commuting operators* $\mathbf{L}^2$ and L_z . They are *not* eigenfunctions of L_x or L_y .

Table of Spherical Harmonics

$Y_{0,0} = \frac{1}{\sqrt{4\pi}}$
$Y_{1,1} = -\sqrt{\frac{3}{8\pi}} e^{i\varphi} \sin \theta$
$Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta$
$Y_{2,2} = \sqrt{\frac{15}{32\pi}} e^{2i\varphi} \sin^2 \theta$
$Y_{2,1} = -\sqrt{\frac{15}{8\pi}} e^{i\varphi} \sin \theta \cos \theta$
$Y_{2,0} = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1)$
$Y_{3,3} = -\sqrt{\frac{35}{64\pi}} e^{3i\varphi} \sin^3 \theta$
$Y_{3,2} = \sqrt{\frac{105}{64\pi}} e^{2i\varphi} \sin^2 \theta \cos \theta$
$Y_{3,1} = -\sqrt{\frac{21}{64\pi}} e^{i\varphi} \sin \theta (5 \cos^2 \theta - 1)$
$Y_{3,0} = \sqrt{\frac{7}{16\pi}} (5 \cos^3 \theta - 3 \cos \theta)$

EXAMPLE 7-1

Use the raising/lowering operators to obtain the (unnormalized) eigenfunctions $Y_{lm}(\theta, \varphi)$ for $l = 2$ for $m = 2, 1, 0$.

SOLUTION We start with

$$Y_{22}(\theta, \varphi) = A \sin^2 \theta e^{2i\varphi}$$

To get $Y_{21}(\theta, \varphi)$ we apply the L_- operator once. From (7-36) we find that

$$Y_{21}(\theta, \varphi) \propto L_- Y_{22}(\theta, \varphi) = \hbar e^{-i\varphi} \left(-\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right) (\sin^2 \theta e^{2i\varphi})$$

A little algebra yields

$$Y_{21}(\theta, \varphi) = B e^{i\varphi} \sin \theta \cos \theta$$

Repeating the application of L_- , we get

$$\begin{aligned} Y_{20}(\theta, \varphi) \propto L_- Y_{21}(\theta, \varphi) &= \hbar e^{-i\varphi} \left(-\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \varphi} \right) B e^{i\varphi} \sin \theta \cos \theta \\ &= B \hbar (-\cos^2 \theta + \sin^2 \theta - \cos^2 \theta) = C(3 \cos^2 \theta - 1) \end{aligned}$$

so that

$$Y_{20}(\theta, \varphi) = C(3 \cos^2 \theta - 1)$$

We conclude with a comment on the connection with classical angular momentum. We expect that for large l the correspondence principle will lead us to the classical limit—that is, we will find something that we can associate with an angular momentum *vector*. We note that when $m = l$ (or $m = -l$) then

$$|Y_{l,\pm l}(\theta, \varphi)|^2 \propto (\sin \theta)^{2l} \quad (7-49)$$

This implies that for large l the probability distribution is strongly confined to the equatorial region around $\theta = \pi/2$. This is as expected: For $m = l$, L_z has its largest value so that L^2 is as close as it can get to L_z^2 . In the limit of large l , the classical limit,

$$\frac{L^2 - L_z^2}{L^2} = \frac{l}{l(l+1)} = \frac{1}{l+1} \rightarrow 0 \quad (7-50)$$

This means that it is possible to line up the angular momentum in a particular direction, here the z -axis. Such an alignment would require

$$\langle L_x^2 \rangle = \langle L_y^2 \rangle = 0$$

which is not possible when quantum mechanical effects become important, and commutation relations cannot be ignored.

7-4 COMMENTS ON THE EXPANSION THEOREM

The general statement of the expansion theorem allows us to expand a state vector in terms of a complete set of states. Consider

$$|\psi\rangle = \sum_{l=0}^{\infty} \sum_{m=-l}^l C_{lm} |l, m\rangle \quad (7-51)$$

With the help of the orthonormality relation

$$\langle l', m' | l, m \rangle = \delta_{ll'} \delta_{mm'} \quad (7-52)$$

we get $C_{lm} = \langle l, m | \psi \rangle$, so that

$$|\psi\rangle = \sum_{l=0}^{\infty} \sum_{m=-l}^l |l, m\rangle \langle l, m | \psi \rangle \quad (7-53)$$

Since this is true for every $|\psi\rangle$, we get the expected completeness relation

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l |l, m\rangle \langle l, m| = \mathbf{1} \quad (7-54)$$

where $\mathbf{1}$ is the unit operator.

If we multiply this on the left by $\langle \theta', \varphi' |$ and on the right by $|\theta, \varphi\rangle$, we get

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l \langle \theta', \varphi' | l, m \rangle \langle l, m | \theta, \varphi \rangle = \langle \theta', \varphi' | \theta, \varphi \rangle$$

which can be written as

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}^*(\theta', \varphi') Y_{lm}(\theta, \varphi) = \frac{1}{\sin \theta} \delta(\theta - \theta') \delta(\varphi - \varphi') \quad (7-55)$$

We can use the expansion theorem to answer the question often raised by students. What is so special about the z -direction? Cannot we align the angular momentum (as far as pos-

sible) with the x -axis? The answer is that this is indeed possible. Such a state, which would be confined near the equatorial plane about the x -axis (in the vicinity of $\varphi = \pi/2$) will be a particular linear combination of the $Y_{lm}(\theta, \varphi)$. Its physical properties will be exactly the same as the Y_{ll} state.

PROBLEMS

- Consider the molecule CN, which may be described by a dumbbell consisting of two masses M_1 and M_2 attached by a rigid rod of length a . The dumbbell rotates *in a plane* about an axis going through the center of mass and perpendicular to it.
 - Write down a Hamiltonian that describes the motion.
 - What is the energy spectrum?
 - Write down an expression for the difference in energy between the ground state and the first excited state in terms of the masses and a .
- Express the spherical harmonics for $l = 0, 1, 2$ in terms of x, y, z .
- Calculate $\langle l, m_1 | L_x | l, m_2 \rangle$ and $\langle l, m_1 | L_y | l, m_2 \rangle$.
- Calculate $\langle l, m_1 | L_x^2 | l, m_2 \rangle$ and $\langle l, m_1 | L_y^2 | l, m_2 \rangle$. (*Hint:* Things will be easier with the use of $L_{\pm}$, $L_{\pm}^2$, and so on.)
- The Hamiltonian for an axially symmetric rotator is

$$H = \frac{L_x^2 + L_y^2}{2I_1} + \frac{L_z^2}{2I_3}$$

- What are the eigenvalues of H ?
 - Sketch the spectrum, assuming that $I_1 > I_3$.
 - What is the spectrum in the limit that I_1 is much larger than I_3 ?
- Use lowering operators to calculate the angular dependence (without worrying about normalization) of $Y_{4m}(\theta, \varphi)$ for $m = 3, 2, 1, 0$. You are given $Y_{44}(\theta, \varphi) = Ae^{4i\varphi} \sin^4 \theta$.
 - A system is described by the Hamiltonian

$$H = \frac{\mathbf{L}^2}{2I} + \alpha L_z$$

What is the energy spectrum of the system?

- Calculate the commutators $[x, L_x]$, $[y, L_x]$, $[z, L_x]$, $[x, L_y]$, $[y, L_y]$, $[z, L_y]$. Do you detect a pattern that will allow you to state the commutators of x, y, z with L_z ?
- Repeat the calculation with $x, \dots$ replaced by $p_x, \dots$ in problem 8.
- Consider a state of total angular momentum $l = 2$. What are the eigenvalues of the operators (a) L_z , (b) $\frac{3}{5}L_x - \frac{4}{5}L_y$, and (c) $2L_x - 6L_y + 3L_z$?
- A particle in a spherically symmetric potential is in a state described by the wave packet

$$\psi(x, y, z) = C(xy + yz + zx)e^{-\alpha r^2}$$

What is the probability that a measurement of the square of the angular momentum yields 0? What is the probability that it yields $6\hbar^2$? If the value of l is found to be 2, what are the relative probabilities for $m = 2, 1, 0, -1, -2$?

- Consider the following model of a perfectly smooth cylinder. It is a ring of equally spaced, identical particles, with mass M/N so that the mass of the ring is M and its moment of inertia is MR^2 , with R

the radius of the ring. Calculate the possible values of the angular momentum. Calculate the energy eigenvalues. What is the energy difference between the ground state of zero angular momentum, and the first rotational state? Show that this approaches infinity as $N \rightarrow \infty$. Contrast this with the comparable energy for a “nicked” cylinder, which lacks the symmetry under the rotation through $2\pi/N$ radians. This example implies that it is impossible to set a perfectly smooth cylinder in rotation, which is consistent with the fact that for a perfectly smooth cylinder such a rotation would be unobservable.

Supplement 7-A

Rotational Invariance

In this supplement we show that the assumption of a central potential implies the conservation of angular momentum. We make use of invariance under rotations. The kinetic energy, which involves $\mathbf{p}^2$, is independent of the direction in which $\mathbf{p}$ points. The central potential $V(r)$ is also invariant under rotations. We show that this invariance implies the conservation of angular momentum.

INVARIANCE UNDER ROTATIONS ABOUT THE z -AXIS

Consider the special case of a rotation through an angle θ about the z -axis: With

$$\begin{aligned}x' &= x \cos \theta - y \sin \theta \\y' &= x \sin \theta + y \cos \theta\end{aligned}\tag{7A-1}$$

it is easy to see that

$$r' = (x'^2 + y'^2 + z'^2)^{1/2} = (x^2 + y^2 + z^2)^{1/2} = r\tag{7A-2}$$

and

$$\begin{aligned}\left(\frac{\partial}{\partial x'}\right)^2 + \left(\frac{\partial}{\partial y'}\right)^2 &= \left(\cos \theta \frac{\partial}{\partial x} - \sin \theta \frac{\partial}{\partial y}\right)^2 + \left(\sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y}\right)^2 \\&= \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2\end{aligned}\tag{7A-3}$$

Since the Hamiltonian has an invariance property, we expect a conservation law, as we saw in the case of parity. To identify the operators that commute with H , let us consider an infinitesimal rotation about the z -axis. Keeping terms of order θ only so that

$$\begin{aligned}x' &= x - \theta y \\y' &= y + \theta x\end{aligned}\tag{7A-4}$$

we require that

$$Hu_E(x - \theta y, y + \theta x, z) = Eu_E(x - \theta y, y + \theta x, z)\tag{7A-5}$$

If we expand this to first order in θ and subtract from it

$$Hu_E(x, y, z) = Eu_E(x, y, z)\tag{7A-6}$$

we obtain from the term linear in θ

$$H\left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\right)u_E(x, y, z) = E\left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\right)u_E(x, y, z)\tag{7A-7}$$

The right side of this equation may be written as

$$\left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\right) E u_E(x, y, z) = \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\right) H u_E(x, y, z) \quad (7A-8)$$

If we define

$$L_z = \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\right) = x p_y - y p_x \quad (7A-9)$$

then (7A-7) and (7A-8) together read

$$(H L_z - L_z H) u_E(x, y, z) = 0$$

Since the $u_E(\mathbf{r})$ form a complete set, this implies the operator relation

$$[H, L_z] = 0 \quad (7A-10)$$

holds. L_z is the z-component of the operator

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (7A-11)$$

which is the angular momentum. Had we taken rotations about the x- and y-axis, we would have found, in addition, that

$$\begin{aligned} [H, L_x] &= 0 \\ [H, L_y] &= 0 \end{aligned} \quad (7A-12)$$

Thus the three components of the angular momentum operators commute with the Hamiltonian; that is, the angular momentum is a constant of the motion. This parallels the classical result that central forces imply conservation of the angular momentum.

Supplement 7-B

Angular Momentum in Spherical Coordinates

We start with spherical coordinates, as defined in Fig. 7-2. We have

$$\begin{aligned}x &= r \sin \theta \cos \varphi \\y &= r \sin \theta \sin \varphi \\z &= r \cos \theta\end{aligned}\tag{7B-1}$$

From this it follows that

$$\begin{aligned}dx &= \sin \theta \cos \varphi dr + r \cos \theta \cos \varphi d\theta - r \sin \theta \sin \varphi d\varphi \\dy &= \sin \theta \sin \varphi dr + r \cos \theta \sin \varphi d\theta + r \sin \theta \cos \varphi d\varphi \\dz &= \cos \theta dr - r \sin \theta d\theta\end{aligned}\tag{7B-2}$$

These may be solved to give

$$\begin{aligned}dr &= \sin \theta \cos \varphi dx + \sin \theta \sin \varphi dy + \cos \theta dz \\d\theta &= \frac{1}{r} (\cos \theta \cos \varphi dx + \cos \theta \sin \varphi dy - \sin \theta dz) \\d\varphi &= \frac{1}{r \sin \theta} (-\sin \varphi dx + \cos \varphi dy)\end{aligned}\tag{7B-3}$$

With the help of these we obtain

$$\begin{aligned}\frac{\partial}{\partial x} &= \frac{\partial r}{\partial x} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x} \frac{\partial}{\partial \theta} + \frac{\partial \varphi}{\partial x} \frac{\partial}{\partial \varphi} \\&= \sin \theta \cos \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \cos \varphi \frac{\partial}{\partial \theta} - \frac{\sin \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi} \\\frac{\partial}{\partial y} &= \sin \theta \sin \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \sin \varphi \frac{\partial}{\partial \theta} + \frac{\cos \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi} \\\frac{\partial}{\partial z} &= \cos \theta \frac{\partial}{\partial r} - \frac{1}{r} \sin \theta \frac{\partial}{\partial \theta}\end{aligned}\tag{7B-4}$$

We may use these to calculate the angular momentum operators in terms of the spherical angles. We have

$$\begin{aligned}\mathbf{L} &= \mathbf{r} \times \mathbf{p}_{\text{op}} = \frac{\hbar}{i} (\mathbf{r} \times \nabla) \\L_z &= \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)\end{aligned}\tag{7B-5}$$

$$\begin{aligned}
&= \frac{\hbar}{i} \left[r \sin \theta \cos \varphi \left(\sin \theta \sin \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \sin \varphi \frac{\partial}{\partial \theta} + \frac{\cos \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi} \right) \right. \\
&\quad \left. - r \sin \theta \sin \varphi \left(\sin \theta \cos \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \cos \varphi \frac{\partial}{\partial \theta} - \frac{\sin \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi} \right) \right] \\
&= \frac{\hbar}{i} \frac{\partial}{\partial \varphi}
\end{aligned} \tag{7B-6}$$

Similarly, we construct

$$\begin{aligned}
L_x &= \frac{\hbar}{i} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \\
&= \frac{\hbar}{i} \left[r \sin \theta \sin \varphi \left(\cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) \right. \\
&\quad \left. - r \cos \theta \left(\sin \theta \sin \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \sin \varphi \frac{\partial}{\partial \theta} + \frac{\cos \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi} \right) \right] \\
&= \frac{\hbar}{i} \left(-\sin \varphi \frac{\partial}{\partial \theta} - \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right)
\end{aligned} \tag{7B-7}$$

and

$$\begin{aligned}
L_y &= \frac{\hbar}{i} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \\
&= \frac{\hbar}{i} \left[r \cos \theta \left(\sin \theta \cos \varphi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \cos \varphi \frac{\partial}{\partial \theta} - \frac{\sin \varphi}{r \sin \theta} \frac{\partial}{\partial \varphi} \right) \right. \\
&\quad \left. - r \sin \theta \cos \varphi \left(\cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) \right] \\
&= \frac{\hbar}{i} \left(\cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right)
\end{aligned} \tag{7B-8}$$

It is fairly straightforward to calculate

$$\begin{aligned}
\mathbf{L}^2 &= L_x^2 + L_y^2 + L_z^2 \\
&= \hbar^2 \left[\left(-\sin \varphi \frac{\partial}{\partial \theta} - \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right) \left(-\sin \varphi \frac{\partial}{\partial \theta} - \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right) \right. \\
&\quad \left. + \left(\cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right) \left(\cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right) + \frac{\partial^2}{\partial \varphi^2} \right]
\end{aligned}$$

We leave it to the reader to do the algebra. The final result is

$$\mathbf{L}^2 = -\hbar^2 \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right) \tag{7B-9}$$

Which is just the expression in eq. (7-13).

The equation

$$\mathbf{L}^2 Y(\theta, \varphi) = \hbar^2 \lambda Y(\theta, \varphi)$$

when written out in spherical coordinates, is

$$\left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} + \lambda \right) Y(\theta, \varphi) = 0 \quad (7B-10)$$

Let us separate variables again. If we write the solution in the form

$$Y(\theta, \varphi) = P(\theta)\Phi(\varphi) \quad (7B-11)$$

We multiply everything by $\sin^2 \theta$, and divide by $P(\theta)\Phi(\varphi)$. This leads to

$$\frac{1}{P(\theta)} \left(\sin^2 \theta \frac{d^2}{d\theta^2} + \sin \theta \cos \theta \frac{d}{d\theta} + \lambda \sin^2 \theta \right) P(\theta) = -\frac{1}{\Phi(\varphi)} \frac{d^2 \Phi(\varphi)}{d\varphi^2}$$

The two sides of the equation depend on different variables. They must therefore each be constant. We call the constant m^2 , without specifying whether this quantity is real or complex. The solution of

$$\frac{\partial^2}{\partial \varphi^2} \Phi(\varphi) = -m^2 \Phi(\varphi) \quad (7B-12)$$

is

$$\Phi(\varphi) = C e^{\pm i m \varphi} \quad (7B-13)$$

The requirement that the solution is single valued—that is, it does not change when $\varphi \rightarrow \varphi + 2\pi$ —requires that m be an integer. When this is substituted into our differential equation, we end up with

$$\left(\frac{d^2}{d\theta^2} + \cot \theta \frac{d}{d\theta} + \lambda \right) P(\theta) = \frac{m^2}{\sin^2 \theta} P(\theta) \quad (7B-14)$$

We now define

$$z = \cos \theta \quad (7B-15)$$

Using

$$\begin{aligned} \frac{d}{d\theta} &= \frac{dz}{d\theta} \frac{d}{dz} = -\sin \theta \frac{d}{dz} \\ \frac{d^2}{d\theta^2} &= \frac{d}{d\theta} \left(-\sin \theta \frac{d}{dz} \right) = -\cos \theta \frac{d}{dz} + \sin^2 \theta \frac{d^2}{dz^2} \end{aligned}$$

and using the fact that $\sin^2 \theta = 1 - z^2$, can rewrite the equation as follows:

$$\left[(1 - z^2) \frac{d^2}{dz^2} - 2z \frac{d}{dz} - \frac{m^2}{1 - z^2} + \lambda \right] P(z) = 0 \quad (7B-16)$$

Consider first the $m = 0$ case. Let us write the solution as

$$P(z) = \sum_{n=0}^{\infty} a_n z^n \quad (7B-17)$$

Some simple manipulations show that the coefficients a_n obey the recurrence relation

$$a_{n+2} = \frac{n^2 + n - \lambda}{(n+1)(n+2)} a_n \quad (7B-18)$$

The series will not terminate if λ is not an integer. In that case, for large n ,

$$\frac{a_{n+2}}{a_n} \rightarrow 1 \quad (7B-19)$$

This means that for some large value of $n = N$ the series approaches a polynomial in z plus

$$1 + z + z^2 + z^3 + \cdots = \frac{1}{1 - z} \quad (7B-20)$$

This, however, is singular at $z = 1$. The only way to evade this singularity is to choose the numerator in (7B-18) to become zero when n reaches some integral value—say, $n = l$. This means that the eigenvalue is

$$\lambda = l(l + 1) \quad (7B-21)$$

and that $P(z)$ is a *polynomial of order l* in the variable z . We will label the polynomial as $P_l(z)$. These polynomials are known as *Legendre* polynomials. A short list follows:

$$\begin{aligned} P_0(z) &= 1 \\ P_1(z) &= z \\ P_2(z) &= \frac{1}{2}(3z^2 - 1) \\ P_3(z) &= \frac{1}{2}(5z^3 - 3z) \\ P_4(z) &= \frac{1}{8}(35z^4 - 30z^2 + 3) \end{aligned} \quad (7B-22)$$

We next observe that the $m \neq 0$ solutions are related to the $m = 0$ solutions. Let us first write the solution of (7B-16) as $P_l^m(z)$. The equation for the $P_l^m(z)$ may be written in the form

$$\frac{d}{dz} \left[(1 - z^2) \frac{dP_l^m(z)}{dz} \right] + \left[l(l + 1) - \frac{m^2}{1 - z^2} \right] P_l^m(z) = 0 \quad (7B-23)$$

If one writes

$$P_l^m(z) = (1 - z^2)^{m/2} F(z) \quad (7B-24)$$

then

$$\frac{dP_l^m(z)}{dz} = -mz(1 - z^2)^{m/2-1} F + (1 - z^2) \frac{dF}{dz}$$

and then

$$\begin{aligned} \frac{d}{dz} \left[(1 - z^2) \frac{dP_l^m(z)}{dz} \right] &= (1 - z^2)^{m/2+1} \frac{d^2 F}{dz^2} - 2z(m + 1)(1 - z^2)^{m/2} \frac{dF}{dz} \\ &\quad + [m(m + 1)z^2 - 1](1 - z^2)^{m/2-1} F \end{aligned}$$

This is to be set equal to

$$\left[\frac{m^2}{1 - z^2} - l(l + 1) \right] (1 - z^2)^{m/2} F$$

After this is done, some rearrangements are made, and the resulting equation is multiplied by $(1 - z^2)^{-m/2}$, one finally obtains the equation

$$(1 - z^2) \frac{d^2 F}{dz^2} - 2z(m + 1) \frac{dF}{dz} + (l - m)(l + m + 1)F = 0 \quad (7B-25)$$

W-34 Supplement 7B Angular Momentum in Spherical Coordinates

One can now show that this equation is satisfied by

$$F(z) = \left(\frac{d}{dz} \right)^m P_l(z) \quad (7B-26)$$

We have therefore shown that

$$P_l^m(z) = (1 - z^2)^{m/2} \left(\frac{d}{dz} \right)^m P_l(z) \quad (7B-27)$$

These functions are known as *associated Legendre functions*.

Note that only m^2 appears in the equation, so that for m in the last equation we can equally well write $|m|$. It is also clear from the form that $|m| \leq l$, since the highest power in $P_l(z)$ is z^l . For m negative we take $(-1)^m P_l^{|m|}(z)$.

Chapter 8

The Schrödinger Equation in Three Dimensions and the Hydrogen Atom

The energy operator H in three dimensions is a simple generalization of the one-dimensional energy. The Schrödinger equation takes the form

$$H\psi(\mathbf{r}) = \left(\frac{\mathbf{p}^2}{2\mu} + V(\mathbf{r}) \right) \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (8-1)$$

Since the letter m (usually reserved for the mass) will be used to denote the z -component of the angular momentum, we use μ for the mass.¹

Since

$$\mathbf{p}^2 = p_x^2 + p_y^2 + p_z^2 \quad (8-2)$$

the generalization of the one-dimensional momentum operator leads to

$$\mathbf{p} = (p_x, p_y, p_z) = \left(\frac{\hbar}{i} \frac{\partial}{\partial x}, \frac{\hbar}{i} \frac{\partial}{\partial y}, \frac{\hbar}{i} \frac{\partial}{\partial z} \right) \quad (8-3)$$

These then show that the three-dimensional Schrödinger equation is a *partial differential equation*, which reads

$$-\frac{\hbar^2}{2\mu} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] \psi(x, y, z) + V(x, y, z) \psi(x, y, z) = E\psi(x, y, z) \quad (8-4)$$

We shall primarily deal with the case of a *central potential* in which the potential energy only depends on the distance from the origin; that is, $V = V(r)$. Before doing that we shall consider a particularly simple case, in which the potential energy is a sum of three terms of the form

$$V(x, y, z) = V_1(x) + V_2(y) + V_3(z) \quad (8-5)$$

An example of such a situation arises if we have an electron in an infinite box of dimensions L_1, L_2, L_3 , or if we have a particle subject to harmonic forces in three dimensions, in which case the potential energy has the form $\frac{1}{2}m\omega^2(x^2 + y^2 + z^2)$. In

¹The reader might think that we have things backward, since the mass appears so much more often than the new parameter. However, when two-body problems are studied, what frequently enters into the kinetic energy formula is the *reduced* mass, and this is usually denoted by the letter μ .

these cases, the motion in the x -direction, for example, is unaffected by the motion in the y - or z -directions. As a consequence we expect that the wave function factorizes into functions of x , y , and z , as follows:

$$\psi(x, y, z) = u_1(x)v_2(y)w_3(z) \quad (8-6)$$

When this is substituted into the energy eigenvalue equation we find that

$$\begin{aligned} -\frac{\hbar^2}{2\mu} \frac{d^2 u_1(x)}{dx^2} + V_1(x)u_1(x) &= E_1 u_1(x) \\ -\frac{\hbar^2}{2\mu} \frac{d^2 v_2(y)}{dy^2} + V_2(y)v_2(y) &= E_2 v_2(y) \\ -\frac{\hbar^2}{2\mu} \frac{d^2 w_3(z)}{dz^2} + V_3(z)w_3(z) &= E_3 w_3(z) \end{aligned} \quad (8-7)$$

with

$$E_1 + E_2 + E_3 = E \quad (8-8)$$

We shall have occasion to discuss this problem in detail in Chapter 11. At this point we turn to the important case of

$$V(x, y, z) = V(r) \quad (8-9)$$

8-1 THE CENTRAL POTENTIAL

For a central potential $V(r)$ the natural coordinates are spherical coordinates. The details of the transition from (x, y, z) to (r, θ, φ) are worked out in Supplement 7-B [www.wiley.com/college/gasiorowicz]. Equation (7B-4) allows us to calculate

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (8-10)$$

and we leave it to the dedicated reader to show that in spherical coordinates

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right) \quad (8-11)$$

Comparison with Eq. (7B-9) shows that this may be written in the form

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{\mathbf{L}^2}{\hbar^2 r^2} \quad (8-12)$$

Thus the Schrödinger equation has the form

$$-\frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) \psi(\mathbf{r}) + \frac{\mathbf{L}^2}{2\mu r^2} \psi(\mathbf{r}) + V(r)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (8-13)$$

We now take advantage of the fact that $Y_{lm}(\theta, \varphi)$ are eigenfunctions of $\mathbf{L}^2$, with eigenvalue $\hbar^2 l(l+1)$. Consequently, if we write

$$\psi(\mathbf{r}) = R_{nl}(r)Y_{lm}(\theta, \varphi) \quad (8-14)$$

we get an equation for the *radial* part of the wave function $R(r)$:

$$-\frac{\hbar^2}{2\mu} \left(\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} - \frac{l(l+1)}{r^2} \right) R_{nl}(r) + V(r)R_{nl}(r) = E_{nl}R_{nl}(r) \quad (8-15)$$

Note that the z -component of the angular momentum does not appear in the equation, so that the quantum number m is absent. This is why we used the labeling (n, l) with n referring to the energy. There are three problems that we will consider in this chapter. The most important application is the *hydrogen atom*, which will be discussed in great detail. We will conclude the chapter with a discussion of the free particle in spherical coordinates, and with an application to the spherical infinite well.

8-2 THE HYDROGEN ATOM

The hydrogen atom is the simplest atom, because it contains only one electron. On the atomic scale, the nucleus is a point particle, and the Schrödinger equation becomes a one-particle equation after the center of mass motion is separated out. It will be useful to treat the nucleus as having charge Ze . The attractive Coulomb potential then is

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r} \quad (8-16)$$

and the radial Schrödinger equation (omitting subscripts), is

$$\left[\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} + \frac{2\mu}{\hbar^2} \left(E + \frac{Ze^2}{4\pi\epsilon_0 r} - \frac{\hbar^2 l(l+1)}{2\mu r^2} \right) \right] R(r) = 0 \quad (8-17)$$

We will concentrate on the bound states—that is, solutions with $E < 0$. It is convenient to introduce dimensionless variables. We define ρ by

$$\rho = \sqrt{\frac{8\mu|E|}{\hbar^2}} r \quad (8-18)$$

In terms of this variable, the equation reads

$$\frac{d^2 R(\rho)}{d\rho^2} + \frac{2}{\rho} \frac{dR(\rho)}{d\rho} - \frac{l(l+1)}{\rho^2} R(\rho) + \left(\frac{\lambda}{\rho} - \frac{1}{4} \right) R(\rho) = 0 \quad (8-19)$$

The dimensionless parameter is

$$\lambda = \frac{Ze^2}{4\pi\epsilon_0 \hbar} \sqrt{\frac{\mu}{2|E|}} \quad (8-20)$$

It can also be written in the form

$$\lambda = Z\alpha \sqrt{\frac{\mu c^2}{2|E|}} \quad (8-21)$$

The second form is easier for computation, since α is known to have the value $1/137$, and the energy is expressed in units of the reduced rest mass energy. On the other hand, (8-20) shows clearly that the velocity of light c does not really appear in the equation, which must be so since the equation is truly nonrelativistic.

8-3 THE ENERGY SPECTRUM

We approach the problem of solving (8-19) in a familiar way. First, we extract the large ρ behavior. For large ρ the potential energy and all the other terms involving inverses of ρ go to zero. What remains is the equation

$$\frac{d^2 R}{d\rho^2} - \frac{1}{4} R = 0 \quad (8-22)$$

The solution, which behaves properly at infinity, is

$$R \approx e^{-\rho/2} \quad (8-23)$$

We now introduce a new function $G(\rho)$ by

$$R(\rho) = e^{-\rho/2} G(\rho) \quad (8-24)$$

We substitute this into (8-19) and do a little algebra to get the equation for $G(\rho)$. This equation is

$$\frac{d^2 G}{d\rho^2} - \left(1 - \frac{2}{\rho}\right) \frac{dG}{d\rho} + \left[\frac{\lambda - 1}{\rho} - \frac{l(l+1)}{\rho^2}\right] G = 0 \quad (8-25)$$

We can take one more step. Consider the equation for very small ρ . In that case

$$\frac{d^2 G}{d\rho^2} + \frac{2}{\rho} \frac{dG}{d\rho} - \frac{l(l+1)}{\rho^2} G \cong 0$$

and this is easily seen to lead to $G(\rho) \propto \rho^l$. If we therefore write

$$G(\rho) = \rho^l H(\rho) \quad (8-26)$$

we obtain an equation for $H(\rho)$. This equation is easily worked out and has the form

$$\frac{d^2 H}{d\rho^2} + \left(\frac{2l+2}{\rho} - 1\right) \frac{dH}{d\rho} + \frac{\lambda - l - 1}{\rho} H = 0 \quad (8-27)$$

Let us now write

$$H(\rho) = \sum_{k=0}^{\infty} a_k \rho^k \quad (8-28)$$

When this is substituted in (8-27) we get

$$\sum_{k=0}^{\infty} a_k \left[k(k-1) \rho^{k-2} + k \left(\frac{2l+2}{\rho} - 1 \right) \rho^{k-1} + (\lambda - l - 1) \rho^{k-1} \right] = 0$$

This can be rearranged to give

$$\sum_{k=0}^{\infty} \rho^{k-1} [(k+1)(k+2l+2)a_{k+1} + (\lambda - l - 1 - k)a_k] = 0$$

Since this must vanish term by term, we must set the coefficient of ρ^{k-1} to zero. This leads to a *recursion relation*, which reads

$$\frac{a_{k+1}}{a_k} = \frac{k + l + 1 - \lambda}{(k + 2l + 2)(k + 1)} \quad (8-29)$$

For large k we get

$$\frac{a_{k+1}}{a_k} \rightarrow \frac{1}{k} \quad (8-30)$$

If the series in (8-28) does not terminate, then for values of k larger than some very large number N , the above relation implies that

$$a_{k+1} \approx \frac{1}{k!} \quad (8-31)$$

This means that the series in (8-28) has the form

$$H(\rho) = (\text{polynomial in } \rho) + e^\rho$$

This behavior is not acceptable, because it implies that $R(\rho)$ grows like $e^{\rho/2}$. To get a square-integrable function we *must require the series to terminate*. This means that for a given l , there will be some integer k , which we denote by n_r for which the series terminates; that is,

$$\lambda = n_r + l + 1 \quad (8-32)$$

Let us introduce the *principal quantum number* n defined by

$$n = n_r + l + 1 \quad (8-33)$$

Then it follows from the fact that $n_r \geq 0$ that

1. $n \geq l + 1$
2. n is an integer.
3. the relation $\lambda = n$ implies that

$$E = -\frac{1}{2} \mu c^2 \frac{(Z\alpha)^2}{n^2} \quad (8-34)$$

This is just the answer that came out of the Bohr model, but there is a difference. We have the wave functions, and we can therefore map out the probability distribution for the electron, and not rely on classical orbits.

The appearance of the reduced mass μ , which is known to be related to the masses of the electron m_e and that of the nucleus M by

$$\mu = \frac{m_e M}{m_e + M} \quad (8-35)$$

means that the frequencies of the light emitted in transitions differs slightly for different hydrogenlike atoms:

$$\omega_{ij} = \frac{E_i - E_j}{\hbar} = \frac{(Z\alpha)^2}{2\hbar} \frac{m_e c^2}{1 + m_e/M} \left(\frac{1}{n_j^2} - \frac{1}{n_i^2} \right) \quad (8-36)$$

The difference between the spectra of hydrogen (a proton) and deuterium (a bound state of a proton and a neutron) led to the discovery of deuterium by Urey and collaborators in 1932. The nuclear mass for deuterium is effectively twice that of the proton.

The Degeneracy of the Spectrum

The energy levels are characterized by n , but our derivation shows that there are really two quantum numbers that appear in the formula, since

$$n = n_r + l + 1 \quad (8-37)$$

Consider the ground state, characterized by $n = 1$, which implies that $n_r = l = 0$. The ground state is unique. Let us next look at the first excited state with $n = 2$. We have two possibilities here: $n_r = 1, l = 0$ and $n_r = 0, l = 1$. It would appear that there are two possible ways of getting the same energy—that is, two different wave functions that lead to the same energy. In fact, for a given eigenvalue l associated with the square of the angular momentum L^2 , there are actually $(2l + 1)$ different eigenfunctions. These are characterized by a quantum number m , appearing in the angular part of the wave function, $Y_{lm}(\theta, \phi)$. We

thus see that for $l = 0$, there is only one state, $m = 0$. For $l = 1$, there are three states with $m = -1, 0, 1$. For $l = 2$, there are five states with $m = -2, -1, 0, 1, 2$. *In general there are $2l + 1$ states for a given value of l .* The Hamiltonian for any central potential only depends on l , and thus *all* the energy eigenvalues will have the degeneracy $(2l + 1)$.

If we now return to $n = 2$, there is *one* state with $n_r = 1, l = 0$, but there are *three* states with $n_r = 0, l = 1$. Thus altogether there are *four* states with energy corresponding to $n = 2$. The wave functions are different. For $n_r = 1, l = 0$, the angular wave function is a constant Y_{00} . For the radial wave function we have the recurrence relation

$$\frac{a_{k+1}}{a_k} = \frac{k - n_r}{(k + 1)(k + 2l + 2)} = -\frac{1}{2} \quad (8-38)$$

so that, with $a_0 = 1$,

$$H(\rho) = 1 - \frac{\rho}{2} \quad (8-39)$$

For the other case we have $n_r = 0$, so that the radial wave function has $H(\rho) = 1$ and there are three angular wave functions $Y_{lm}(\theta, \phi)$ with $m = -1, 0, 1$. The shapes of the absolute squares of these functions will be shown in Fig. 8-4.

For $n = 3$, we have the following possibilities: $n_r = 2, l = 0$ (one state), $n_r = 1, l = 1$ (three states), and $n_r = 0, l = 2$ (five states). Thus the total number of states is $1 + 3 + 5 = 9$. The wave functions for these states are easily obtained using the recursion relation. For $n_r = 2, l = 0$, we have $a_2/a_1 = -1/6$ and $a_1/a_0 = -1$ so that

$$H(\rho) = 1 - \rho + \frac{\rho^2}{6} \quad (8-40)$$

while the angular wave function is constant. For $n_r = 1, l = 1$ the radial function will be

$$H(\rho) = 1 - \frac{\rho}{4} \quad (8-41)$$

associated with three angular wave functions Y_{lm} , and so on. Quite generally the degeneracy for a general n is

$$\sum_{l=0}^{l_{\max}} (2l + 1) = \sum_{l=0}^{n-1} (2l + 1) = n^2 \quad (8-42)$$

Figure 8-1 shows the spectrum.

As mentioned before, we expect a $(2l + 1)$ degeneracy for a radial potential, because the radial Hamiltonian only depends on L^2 . Here there is a larger degeneracy, which is peculiar to a $1/r$ potential. We can easily check that a modification of the potential removes the degeneracy. Suppose we modify the Coulomb potential by a term that changes the potential for small r . We choose a convenient modification:

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r} + \frac{\hbar^2 g^2}{2\mu r^2} \quad (8-43)$$

With this potential the radial equation is unchanged, except that where we had $l(l + 1)/r^2$ we now have $l^*(l^* + 1)/r^2$ where $l^*(l^* + 1) = l(l + 1) + g^2$, that is $l^* = -1/2 + \sqrt{(l + 1/2)^2 + g^2}$. This gives the same energy spectrum as (8-34) with (8-33) in which l is replaced by l^* , so that

$$E = -\frac{1}{2} \mu c^2 \frac{(Z\alpha)^2}{[n_r + 1/2 + \sqrt{(l + 1/2)^2 + g^2}]^2} \quad (8-44)$$

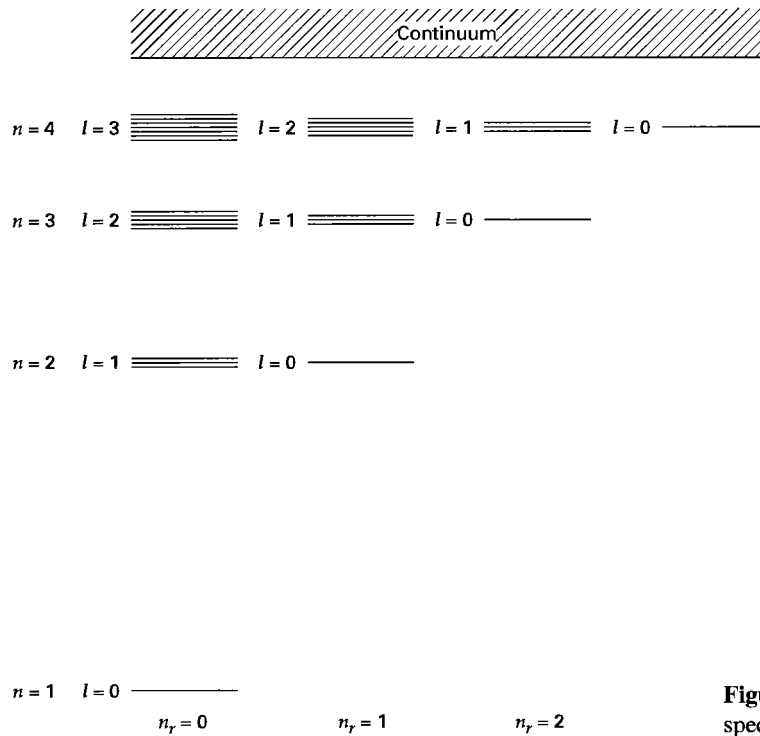


Figure 8-1 The hydrogen spectrum (not to scale).

Here we see, for example, that the energies for $(n_r = 1, l = 2)$ and $(n_r = 2, l = 1)$ are no longer degenerate. Degeneracy usually arises from the existence of an additional constant of the motion. There exists, indeed, an additional operator that commutes with H , L^2 , and L_z , for the special case of a $1/r$ potential.²

In classical mechanics the $1/r$ potential is distinguished by the fact that the elliptical orbits maintain their orientation in space, instead of precessing (Fig. 8-2). The precession of the planetary orbits around the sun are well understood as resulting from the gravitational effects of other planets. After all of these are taken into account, the remaining precession effects can be explained by Einstein's theory of gravitation, which modifies Newton's $1/r$ potential. The largest of these effects is the precession of the perihelion of Mercury by 42 seconds of arc per century, and its explanation by Einstein in 1915 played an enormous role in the rapid acceptance of the general theory of relativity.

²It can be shown that there is one more operator that commutes with H , L^2 , and L_z . It is the so-called *Lenz vector*, whose form is

$$\mathbf{A} = \frac{1}{2\mu\alpha} [\mathbf{L} \times \mathbf{p} - \mathbf{p} \times \mathbf{L}] + \frac{\mathbf{r}}{r}$$

or, equivalently,

$$A_i = \frac{1}{2\mu\alpha} [-x_i p^2 - p^2 x_i + 2(\mathbf{r} \cdot \mathbf{p}) p_i - 4i\hbar p_{il}] + \frac{x_i}{r}$$

W. Pauli used this to calculate the hydrogen spectrum using the same operator methods that we used in our discussion of the harmonic oscillator. This was done shortly before the discovery of the Schrödinger equation, and showed that the Heisenberg, Born, and Jordan formulation of quantum mechanics was correct.

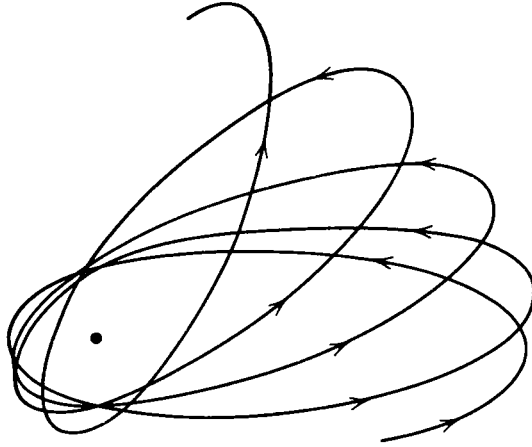


Figure 8-2 Orbits for a potential that does not have the exact $1/r$ form do not close upon themselves and precess as shown here. The orbits remain planar as long as the potential is radial.

The function $H(\rho)$ is a known function, the so-called *associated Laguerre polynomial*. It is

$$H(\rho) = L_{n-l-1}^{(2l+1)}(\rho) \quad (8-45)$$

and its general form can be found in any book dealing with *special functions*. The series expansion is given by

$$L_n^\alpha(\rho) = \sum_{m=0}^n \binom{n+\alpha}{n-m} \frac{(-\rho)^m}{m!} \quad (8-46)$$

When all of this is put together, we can find the radial functions expressed in terms of r . In what follows, $a_0 = \hbar/\mu\alpha$. This is the *Bohr radius* of the ground state. The following qualitative features emerge from the sampling of eigensolutions:

Table of Radial Functions

$R_{10}(r) = 2 \left(\frac{Z}{a_0} \right)^{3/2} e^{-Zr/a_0}$
$R_{20}(r) = 2 \left(\frac{Z}{2a_0} \right)^{3/2} \left(1 - \frac{Zr}{2a_0} \right) e^{-Zr/2a_0}$
$R_{21}(r) = \frac{1}{\sqrt{3}} \left(\frac{Z}{2a_0} \right)^{3/2} \frac{Zr}{a_0} e^{-Zr/2a_0}$
$R_{30}(r) = 2 \left(\frac{Z}{3a_0} \right)^{3/2} \left[1 - \frac{2Zr}{3a_0} + \frac{2(Zr)^2}{27a_0^2} \right] e^{-Zr/3a_0}$
$R_{31}(r) = \frac{4\sqrt{2}}{3} \left(\frac{Z}{3a_0} \right)^{3/2} \frac{Zr}{a_0} \left(1 - \frac{Zr}{6a_0} \right) e^{-Zr/3a_0}$
$R_{32}(r) = \frac{2\sqrt{2}}{27\sqrt{5}} \left(\frac{Z}{3a_0} \right)^{3/2} \left(\frac{Zr}{a_0} \right)^2 e^{-Zr/3a_0}$

- (a) The behavior of r^l for small r , which forces the wave function to stay small for a range of radii that increases with l , is a consequence of the centrifugal repulsive barrier that keeps the electrons from coming close to the nucleus.

- (b) The relation (8-46) shows that $H(\rho)$ is a polynomial of degree $n_r = n - l - 1$, and thus it has n_r radial nodes (zeros). There will be $n - l$ “bumps” in the probability density distribution

$$P(r) = r^2 [R_{nl}]^2 \quad (8-47)$$

When, for a given n , l has its largest value $l = n - 1$, then there is only one bump. As the table suggests, and as can be seen from the solution to the differential equation,

$$R_{n,n-1}(r) \propto r^{n-1} e^{-Zr/a_0 n} \quad (8-48)$$

Hence $P(r) \propto r^{2n} e^{-2Zr/a_0 n}$ will peak at a value of r determined by

$$\frac{dP(r)}{dr} = \left(2nr^{2n-1} - \frac{2Z}{a_0 n} r^{2n} \right) e^{-2Zr/a_0 n} = 0 \quad (8-49)$$

that is, at

$$r = \frac{n^2 a_0}{Z} \quad (8-50)$$

which is the Bohr atom value for circular orbits. Smaller values of l give probability distributions with more bumps. One can show that they correspond to elliptical orbits in the large quantum number limit.

- (c) Plots of the radial probability density $P(r)$ for finding the electron in a shell at a distance r from the origin can be constructed with the help of the radial eigenfunctions. Figure 8-3 shows the general pattern of $r^2 R_{nl}(r)$. We must remember that the wave function also has an angular part. Plots of the probability distribution $\sin \theta (P_l^m(\cos \theta))^2$ as a function of θ are given in Fig. 8-4. As m increases, the probability density is seen to shift from the z -axis toward the equatorial plane. As noted before, as $|m| \rightarrow l$, $(P_l^m(\cos \theta))^2 \rightarrow \sin^2 \theta$, and this function peaks about $\theta = \pi/2$. As l increases, the width of the peak can be shown to decrease like $\sqrt{1/l}$, and thus for large quantum numbers we get the classical picture of planar orbits.
- (d) Given the wave functions, we can calculate

$$\langle r^k \rangle = \int_0^\infty dr r^{2+k} [R_{nl}(r)]^2 \quad (8-51)$$

Some useful expectation values are given below:

$$\begin{aligned} \langle r \rangle &= \frac{a_0}{2Z} [3n^2 - l(l+1)] \\ \langle r^2 \rangle &= \frac{a_0^2 n^2}{2Z^2} [5n^2 + 1 - 3l(l+1)] \\ \left\langle \frac{1}{r} \right\rangle &= \frac{Z}{a_0 n^2} \\ \left\langle \frac{1}{r^2} \right\rangle &= \frac{Z^2}{a_0^2 n^3 (l + \frac{1}{2})} \\ \left\langle \frac{1}{r^3} \right\rangle &= \frac{Z^3}{a_0^3 n^3 l (l + \frac{1}{2})(l+1)} \end{aligned} \quad (8-52)$$

Some of these calculations can be bypassed by using a theorem of W. Pauli discussed in Supplement 8-A [www.wiley.com/college/gasiorowicz].

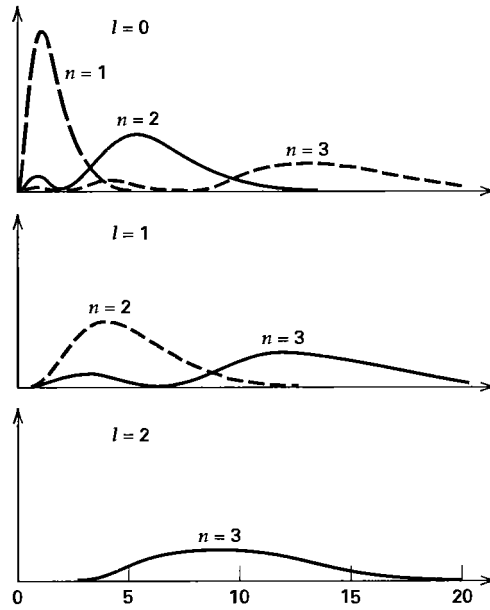


Figure 8-3 Plot of radial probability distributions $r^2(R_{nl}(r))^2$ for $l = 0, 1, 2$. The horizontal axis is r in units of a_0 .

8-4 THE FREE PARTICLE*

In this case $V(r) = 0$, but there is still a centrifugal barrier present. The radial equation (8-15) takes the form

$$\left[\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} - \frac{l(l+1)}{r^2} \right] R(r) + k^2 R(r) = 0 \quad (8-53)$$

If we introduce the variable $\rho = kr$, we get

$$\frac{d^2 R}{d\rho^2} + \frac{2}{\rho} \frac{dR}{d\rho} - \frac{l(l+1)}{\rho^2} R + R = 0 \quad (8-54)$$

or with $u(\rho) = \rho R(\rho)$,

$$\frac{d^2 u}{d\rho^2} - \frac{l(l+1)}{\rho^2} u + u = 0$$

For $l = 0$ the equation for $u(\rho)$ is just $d^2 u/d\rho^2 + u = 0$, so that the solutions are $\sin \rho$ and $\cos \rho$; that is, the regular solution is

$$R_0(\rho) = \frac{\sin \rho}{\rho} \quad (8-55)$$

and the irregular solution is

$$R_0(\rho) = \frac{\cos \rho}{\rho} \quad (8-56)$$

*For particles in a potential, see Supplement 8-B [www.wiley.com/college/gasiorowicz].

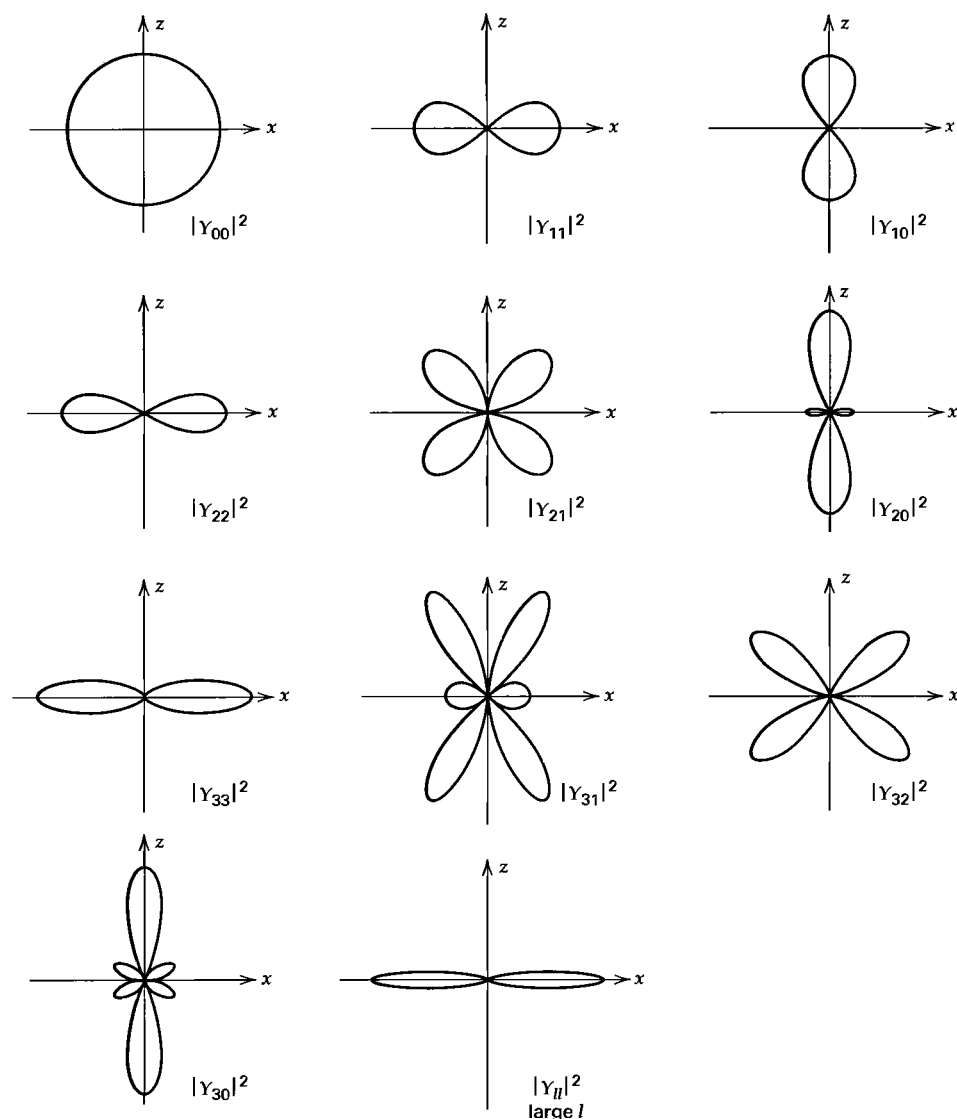


Figure 8-4 Distributions of $|Y_{lm}(\theta, \phi)|^2$. The sketches represent sections of the distributions made in the z - x plane. It should be understood that the three-dimensional distributions are obtained by rotating the figures about the z -axis.

For general l the solutions can be expressed in terms of simple functions. These are known as *spherical Bessel functions*. The regular solution is $j_l(\rho)$, which may be written in the form

$$j_l(\rho) = (-\rho)^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \left(\frac{\sin \rho}{\rho} \right) \quad (8-57)$$

and the irregular one, called the spherical Neumann function $n_l(\rho)$, has the form

$$n_l(\rho) = -(-\rho)^l \left(\frac{1}{\rho} \frac{d}{d\rho} \right)^l \left(\frac{\cos \rho}{\rho} \right) \quad (8-58)$$

The first few functions are listed below.

$$\begin{aligned}
 j_0(\rho) &= \frac{\sin \rho}{\rho} & n_0(\rho) &= -\frac{\cos \rho}{\rho} \\
 j_1(\rho) &= \frac{\sin \rho}{\rho^2} - \frac{\cos \rho}{\rho} & n_1(\rho) &= -\frac{\cos \rho}{\rho^2} - \frac{\sin \rho}{\rho} \\
 j_2(\rho) &= \left(\frac{3}{\rho^3} - \frac{1}{\rho}\right) \sin \rho - \frac{3}{\rho^2} \cos \rho & n_2(\rho) &= -\left(\frac{3}{\rho^3} - \frac{1}{\rho}\right) \cos \rho - \frac{3}{\rho^2} \sin \rho
 \end{aligned} \tag{8-59}$$

The combinations that will be of interest for large ρ are the *spherical Hankel functions*

$$h_l^{(1)}(\rho) = j_l(\rho) + in_l(\rho) \tag{8-60}$$

and

$$h_l^{(2)}(\rho) = [h_l^{(1)}(\rho)]^* \tag{8-61}$$

Again the first few spherical Hankel functions are

$$\begin{aligned}
 h_0^{(1)}(\rho) &= \frac{e^{i\rho}}{i\rho} \\
 h_1^{(1)}(\rho) &= \frac{e^{i\rho}}{\rho} \left(1 + \frac{i}{\rho}\right) \\
 h_2^{(1)}(\rho) &= \frac{i e^{i\rho}}{\rho} \left(1 + \frac{3i}{\rho} - \frac{3}{\rho^2}\right)
 \end{aligned} \tag{8-62}$$

Of special interest are

(a) The behavior near the origin: For $\rho \ll l$, it turns out that

$$j_l(\rho) \approx \frac{\rho^l}{1 \cdot 3 \cdot 5 \cdots (2l+1)} \tag{8-63}$$

and

$$n_l(\rho) \approx \frac{1 \cdot 3 \cdot 5 \cdots (2l-1)}{\rho^{l+1}} \tag{8-64}$$

(b) For $\rho \gg l$, we have the asymptotic expressions

$$j_l(\rho) \approx \frac{1}{\rho} \sin \left(\rho - \frac{l\pi}{2} \right) \tag{8-65}$$

and

$$n_l(\rho) \approx -\frac{1}{\rho} \cos \left(\rho - \frac{l\pi}{2} \right) \tag{8-66}$$

so that

$$h_l^{(1)}(\rho) \approx -\frac{i}{\rho} e^{i(\rho - l\pi/2)} \tag{8-67}$$

The solution that is regular at the origin is

$$R_l(r) = j_l(kr) \tag{8-68}$$

Its asymptotic form is, using (8-65),

$$R_l(r) \approx -\frac{1}{2ikr} [e^{-i(kr - l\pi/2)} - e^{i(kr - l\pi/2)}] \tag{8-69}$$

8-5 PARTICLE IN AN INFINITE SPHERICAL WELL

We conclude this chapter with a discussion of a particle in an infinite spherical well in three dimensions. Here

$$V(r) = \begin{cases} 0 & r \leq a \\ \infty & r > a \end{cases} \quad (8-70)$$

With

$$\frac{2\mu E}{\hbar^2} = k^2 \quad (8-71)$$

the solution regular at $r = 0$ is

$$R(r) = A j_l(kr) \quad (8-72)$$

The eigenvalues are determined by the condition that the solution vanishes at $r = a$ —that is, by

$$j_l(ka) = 0 \quad (8-73)$$

The roots of the spherical Bessel function for a few values of l are given in the following table.

$l = 0$	1	2	3	4	5	6
3.14	4.49	5.76	6.99	8.18	9.36	10.51
6.28	7.73	9.10	10.42	11.70	12.97	
9.42	10.90	12.32				
12.57						

It follows from (8-63) that for ka large (actually $ka \gg l$) the roots are given by

$$ka = \left(n + \frac{1}{2} l \right) \pi \quad (8-73')$$

The eigensolutions are thus

$$u_{nl}(r) = A j_l(k_{nl}r) Y_{lm}(\theta, \varphi) \quad (8-74)$$

The k_{nl} are given by the condition in (8-73). A is a numerical constant required for normalization purposes. Note that

$$\int d^3r u_{n'l'}^*(\mathbf{r}) u_{nl}(\mathbf{r}) = 0 \quad (8-75)$$

when any of the $n \neq n'$, $l \neq l'$. The $Y_{lm}(\theta, \varphi)$ have the property that they are orthogonal for differing values of the quantum numbers that label them, in particular when $l \neq l'$. Orthogonality with respect to the label n , which distinguishes different energy eigenvalues for fixed l , must surely be there; in fact, one finds in discussions of Bessel functions that

$$\int_0^1 t^2 dt j_l(\alpha_m t) j_l(\alpha_n t) = 0 \quad (8-76)$$

for $m \neq n$, where the α_n are defined by $j_l(\alpha_n) = 0$. The relation is equivalent to the orthogonality of the radial functions corresponding to different values of the radial quantum number n .

The spectrum of the infinite square well can be described as follows: If the first root for a given l is labeled $n = 1$, the second root $n = 2$, and so on, and if we use the accepted spectroscopic notation for the l -values

l	0	1	2	3	4	5	6	...
Label	S	P	D	F	G	H	I	...

then the order in which the levels occur is

$$1S, 1P, 1D, 2S, 1F, 2P, 1G, 2D, 1H, 3S, 2F, 1I, 3P, 2G, 3D, 4S, 2H, \dots$$

The ordering of levels in a given potential has some relevance to the *shell structure* that is observed in atoms and in nuclei. This follows from the fact that the degeneracy of states for a particular value of l is $(2l + 1)$, and, as will be discussed in Chapter 11, only two particles of a certain kind (electrons, protons, neutrons) can be assigned to a given energy level.

PROBLEMS

1. Consider a special case of eq. (8-5) in which each of the potentials V_1 , V_2 and V_3 are identical, in each case of the form

$$V(x) = 0 \quad x \leq a$$

$$= \infty \quad x > a$$

and similarly for y and z . Use what you learned about the one-dimensional potential in Chapter 3 to find the eigenvalues and eigenfunctions for a particle in such a box.

2. What is the ground state energy for the potential in Problem 1? List the values of the lowest 10 energy levels and label them by their appropriate quantum numbers. What is the degeneracy of the levels on your list?
3. Consider a potential of the form

$$V(x, y, z) = \frac{1}{2} \mu \omega^2 (x^2 + y^2 + z^2)$$

Using what you know about the one-dimensional simple harmonic oscillator, write down an expression for the energy. What is the energy of the ground state for this potential?

4. List the values of the lowest 10 energy levels in units of $\hbar\omega$, and give the degeneracy corresponding to each of the energy eigenvalues on your list.
5. Work out the Schrödinger equation in polar coordinates ρ, ϕ , with $x = \rho \cos \phi$, $y = \rho \sin \phi$, for a potential that depends only on ρ . If the solution of the equation $\Psi(\rho, \phi)$ is written as $R(\rho)\Phi(\phi)$, what is the equation obeyed by $\Phi(\phi)$? What is the equation for $R(\rho)$?
6. Compare the wavelengths of the $2P \rightarrow 1S$ transitions in (1) hydrogen, (1) deuterium (nuclear mass = $2 \times$ proton mass), (3) positronium (a bound state of an electron and a positron, whose mass is the same as that of an electron).
7. An electron is in the ground state of tritium, for which the nucleus consists of a proton and two neutrons. A nuclear reaction instantaneously changes the nucleus to He^3 —that is, two protons and one neutron. Calculate the probability that the electron remains in the ground state of He^3 .
8. The relativistic analog of the Schrödinger equation for a spin 0 electron (thus not applicable to the real electron) is the operator version of

$$(E - V)^2 = p^2 c^2 + m^2 c^4$$

that is,

$$\left(\frac{E}{\hbar c} + \frac{Ze^2}{4\pi\epsilon_0\hbar c} \frac{1}{r}\right)^2 \psi = -\nabla^2 \psi + \left(\frac{mc}{\hbar}\right)^2 \psi$$

(a) Find the radial equation.

(b) Find the eigenvalue spectrum by noting the close relationship of the radial equation obtained in (a) with the radial equation for the hydrogen atom problem.

9. Using the expression for $\langle 1/r \rangle_{n,l}$ calculate the expression for

$$\langle T \rangle_{n,l} = \left\langle \frac{p^2}{2m} \right\rangle_{n,l}$$

for an arbitrary hydrogen atom eigenstate (with Z arbitrary). Show that generally for this potential

$$\langle T \rangle = -\frac{1}{2} \langle V \rangle$$

This is a special example of the *Virial theorem*.

10. An electron in the Coulomb field of a proton is in a state described by the wave function

$$\frac{1}{6}[4\psi_{100}(\mathbf{r}) + 3\psi_{211}(\mathbf{r}) - \psi_{210}(\mathbf{r}) + \sqrt{10}\psi_{21-1}(\mathbf{r})]$$

(a) What is the expectation value of the energy?

(b) What is the expectation value of $\mathbf{L}^2$?

(c) What is the expectation value of L_z ?

11. An electron in the Coulomb field of a proton is in a state described by the wave function

$$\psi(\mathbf{r}) = \left(\frac{\alpha}{\sqrt{\pi}}\right)^{3/2} e^{-\alpha^2 r^2/2}$$

Write out an expression for the probability that it will be found in the ground state of the hydrogen atom.

12. The expectation value of $f(\mathbf{r}, \mathbf{p})$ in any stationary state is a constant. Calculate

$$0 = \frac{d}{dt} \langle \mathbf{r} \cdot \mathbf{p} \rangle = \frac{i}{\hbar} \langle [H, \mathbf{r} \cdot \mathbf{p}] \rangle$$

for a Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + V(r)$$

and show that

$$\left\langle \frac{\mathbf{p}^2}{m} \right\rangle = \langle \mathbf{r} \cdot \nabla V(r) \rangle$$

Use this to establish the result of Problem 9. Also use this result to calculate $\langle 1/r \rangle$.

13. Use the techniques developed in this chapter to discuss the three-dimensional harmonic oscillator problem, with

$$H = \frac{\mathbf{p}^2}{2m} + \frac{1}{2}m\omega^2 r^2$$

Note that the associated Laguerre polynomials also appear in this problem.

Supplement 8-A

A Useful Theorem

The following useful result appears in Pauli's 1930 "Handbuch Article on Quantum Theory":

Consider eigenvalues and eigenfunctions of a Hamiltonian that depends on some parameter—for example, the mass of the electron, or the charge of the electron, or any other parameter that may appear in more complicated problems. The Schrödinger eigenvalue equation may then be written with the parameter α explicitly indicated as,

$$H(\alpha)u_n(\mathbf{r}, \alpha) = E(\alpha)u_n(\mathbf{r}, \alpha) \quad (8A-1)$$

It follows that with the eigenfunctions normalized to unity,

$$\int d^3r u_n^*(\mathbf{r}, \alpha)u_n(\mathbf{r}, \alpha) = 1 \quad (8A-2)$$

that

$$E(\alpha) = \int d^3r u_n^*(\mathbf{r}, \alpha)H(\alpha)u_n(\mathbf{r}, \alpha) \quad (8A-3)$$

Let us now differentiate both sides with respect to α . We get

$$\begin{aligned} \frac{\partial E(\alpha)}{\partial \alpha} &= \int d^3r \frac{\partial u_n^*(\mathbf{r}, \alpha)}{\partial \alpha} H(\alpha)u_n(\mathbf{r}, \alpha) \\ &\quad + \int d^3r u_n^*(\mathbf{r}, \alpha)H(\alpha) \frac{\partial u_n(\mathbf{r}, \alpha)}{\partial \alpha} + \int d^3r u_n^*(\mathbf{r}, \alpha) \frac{\partial H(\alpha)}{\partial \alpha} u_n(\mathbf{r}, \alpha) \end{aligned}$$

Consider now the first two terms on the right-hand side. Using the eigenvalue equation and its complex conjugate (with hermiticity of H), we see that they add up to

$$\begin{aligned} E(\alpha) \int d^3r \frac{\partial u_n^*(\mathbf{r}, \alpha)}{\partial \alpha} u_n(\mathbf{r}, \alpha) + E(\alpha) \int d^3r u_n^*(\mathbf{r}, \alpha) \frac{\partial u_n(\mathbf{r}, \alpha)}{\partial \alpha} \\ = E(\alpha) \frac{\partial}{\partial \alpha} \int d^3r u_n^*(\mathbf{r}, \alpha)u_n(\mathbf{r}, \alpha) = 0 \end{aligned}$$

We are therefore left with

$$\frac{\partial E(\alpha)}{\partial \alpha} = \int d^3r u_n^*(\mathbf{r}, \alpha) \frac{\partial H(\alpha)}{\partial \alpha} u_n(\mathbf{r}, \alpha) = \left\langle \frac{\partial H(\alpha)}{\partial \alpha} \right\rangle \quad (8A-4)$$

The utility of this result is somewhat limited, because it requires knowing the exact eigenvalues and, for the calculation on the right-hand side, the exact eigenfunctions.¹ Nevertheless, the theorem does allow us certain shortcuts in calculations.

¹The extension of this to certain approximate solutions is due to R. P. Feynman and H. Hellmann. See Problem 10 in Chapter 14.

W-36 Supplement 8-A A Useful Theorem

Consider, for example, the one-dimensional simple harmonic oscillator, for which the Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2 \quad (8A-5)$$

The eigenvalues are known to be

$$E_n = \hbar\omega(n + \frac{1}{2}) \quad (8A-6)$$

If we differentiate E_n with respect to ω , and if we note that

$$\frac{\partial H}{\partial \omega} = m\omega x^2$$

we can immediately make the identification

$$\hbar(n + \frac{1}{2}) = m\omega \langle x^2 \rangle_n$$

or

$$\langle x^2 \rangle_n = \frac{\hbar}{m\omega} \left(n + \frac{1}{2} \right) = \frac{E_n}{m\omega^2} \quad (8A-7)$$

Examples of relevance to the hydrogen atom are of particular interest. In the Hamiltonian, the factor

$$\frac{-e^2}{4\pi\epsilon_0 r} = -\frac{\hbar c \alpha}{r}$$

appears. The eigenvalue has the form

$$E_{nl} = -\frac{1}{2} \frac{mc^2 \alpha^2}{n^2}$$

If we take as our parameter to be α , then we get

$$-\hbar c \left\langle \frac{1}{r} \right\rangle_{n,l} = \frac{\partial}{\partial \alpha} E_{nl} = -\frac{mc^2 \alpha}{n^2} \quad (8A-8)$$

so that

$$\left\langle \frac{1}{r} \right\rangle_{nl} = \frac{mc\alpha}{\hbar n^2} = \frac{1}{a_0 n^2} \quad (8A-9)$$

In the *radial* Hamiltonian, there is a term

$$\frac{\hbar^2}{2m} \frac{l(l+1)}{r^2}$$

If we treat l as the parameter and recall that $n = n_r + l + 1$, we get

$$\frac{\hbar^2}{2m} \left\langle \frac{2l+1}{r^2} \right\rangle = \frac{1}{2} mc^2 \alpha^2 \frac{2}{n^3} \quad (8A-10)$$

which is equivalent to

$$\left\langle \frac{1}{r^2} \right\rangle_{nl} = \frac{1}{a_0^2 n^3 (l + \frac{1}{2})} \quad (8A-11)$$

Using an observation of J. Schwinger that the average force in a stationary state must vanish, we can proceed from

$$\begin{aligned} F &= -\frac{dV(r)}{dr} = -\frac{d}{dr} \left(-\frac{e^2}{4\pi\epsilon_0 r} + \frac{\hbar^2 l(l+1)}{2mr^2} \right) \\ &= -\frac{e^2}{4\pi\epsilon_0 r^2} + \frac{\hbar^2 l(l+1)}{mr^3} \end{aligned} \quad (8A-12)$$

to $\langle F(r) \rangle = 0$ and thus obtain

$$\left\langle \frac{1}{r^3} \right\rangle_{nl} = \frac{m}{\hbar^2 l(l+1)} \frac{e^2}{4\pi\epsilon_0} \left\langle \frac{1}{r^2} \right\rangle_{nl} = \frac{1}{a_0^3 n^3 l(l+\frac{1}{2})(l+1)} \quad (8A-13)$$

Supplement 8-B

The Square Well, Continuum Solutions

We saw in eq. (8-69) that asymptotically the free-particle solution has the form

$$R_l(r) \rightarrow -\frac{1}{2ikr} [e^{-i(kr-l\pi/2)} - e^{i(kr-l\pi/2)}]$$

We now assert that the first term is an incoming spherical wave, and the second is an outgoing spherical wave. The description is arrived at in the following way.

Consider the three-dimensional probability flux

$$\mathbf{j} = \frac{\hbar}{2i\mu} (\psi^*(\mathbf{r})\nabla\psi(\mathbf{r}) - \nabla\psi^*(\mathbf{r})\psi(\mathbf{r}))$$

We shall see that it is only the radial flux that is of interest for large r . The radial flux, integrated over all angles, is thus

$$\int d\Omega \mathbf{i}_r \cdot \mathbf{j}(\mathbf{r}) = \frac{\hbar}{2i\mu} \int d\Omega \left(\psi^* \frac{\partial\psi}{\partial r} - \frac{\partial\psi^*}{\partial r} \psi \right) \quad (8B-1)$$

For a solution of the form

$$\psi(\mathbf{r}) = C \frac{e^{\pm ikr}}{r} Y_{lm}(\theta, \varphi) \quad (8B-2)$$

with $\int d\Omega |Y_{lm}(\theta, \varphi)|^2 = 1$ the right-hand side of (8B-1) can easily be evaluated, and we find that

$$\int d\Omega j_r = \pm \frac{\hbar k |C|^2}{\mu} \frac{1}{r^2} \quad (8B-3)$$

The $\pm$ sign describes outgoing/incoming flux. The factor $1/r^2$ that emerges from our calculation is actually necessary for flux conservation, since the flux going through a spherical surface of radius r is

$$\int r^2 d\Omega j_r = \pm \frac{\hbar k |C|^2}{\mu}, \quad \text{independent of } r$$

We therefore see that for the free-particle solution the incoming flux is equal in magnitude to the outgoing flux, which is what it should be, because there are no sources of flux.

We now note that in the presence of a potential flux is still conserved. Any solution will asymptotically consist of an incoming spherical wave and an outgoing spherical

wave, with the constraint that the magnitude of the incoming flux and the outgoing flux be equal. Thus if the asymptotic solution has the form

$$R_l(kr) \rightarrow -\frac{1}{2ikr} (e^{-i(kr-l\pi/2)} - S_l(k)e^{i(kr-l\pi/2)}) \quad (8B-4)$$

then it is required that

$$|S_l(k)|^2 = 1 \quad (8B-5)$$

We write $S_l(k)$ in the standard form

$$S_l(k) = e^{2i\delta_l(k)} \quad (8B-6)$$

The real function $\delta_l(k)$ is called the *phase shift*, because the asymptotic form of the radial function (8B-4) may be rewritten in the form

$$R_l(r) \rightarrow e^{i\delta_l(k)} \frac{\sin(kr - l\pi/2 + \delta_l(k))}{kr} \quad (8B-7)$$

Aside from the irrelevant phase factor in front, this differs from the asymptotic form of the free-particle solution only by a shift in phase of the argument.

We note parenthetically that with a solution that has a $1/r$ behavior, the flux in any direction other than radial goes to zero as $1/r^2$, and we were therefore justified in only considering the radial flux at large values of r .

Let us now consider the special case of a square well. The above argument shows us that we only need to consider the phase shift, since at large distances from the well the only deviation from free particle behavior is the phase shift.

We again consider the well

$$\begin{aligned} V(r) &= -V_0 & \text{for } r \leq a \\ &= 0 & \text{for } r > a \end{aligned} \quad (8B-8)$$

We again use the notation

$$\kappa^2 = \frac{2\mu(E + V_0)}{\hbar^2} \quad (8B-9)$$

Now the solution for $r \leq a$ must be regular at the origin, so that it has the form

$$R_l(r) = A j_l(\kappa r) \quad r \leq a \quad (8B-10)$$

The solution for $r > a$ will contain an irregular part, so that we have

$$R_l(r) = B j_l(kr) + C n_l(kr) \quad r \geq a \quad (8B-11)$$

The matching of $\frac{1}{R_l(r)} \frac{dR_l(r)}{dr}$ at $r = a$ yields an expression

$$\kappa \left[\frac{dj_l(\rho)/d\rho}{j_l(\rho)} \right]_{\rho=\kappa a} = k \left[\frac{Bdj_l/d\rho + Cdn_l/d\rho}{Bj_l(\rho) + Cn_l(\rho)} \right]_{\rho=ka} \quad (8B-12)$$

from which the ratio C/B can be calculated. The ratio can be related to the phase shift. We do this by looking at the asymptotic form of the larger r solution, which has the form

$$R_l(r) \rightarrow B \frac{\sin(kr - l\pi/2)}{kr} - C \frac{\cos(kr - l\pi/2)}{kr}$$

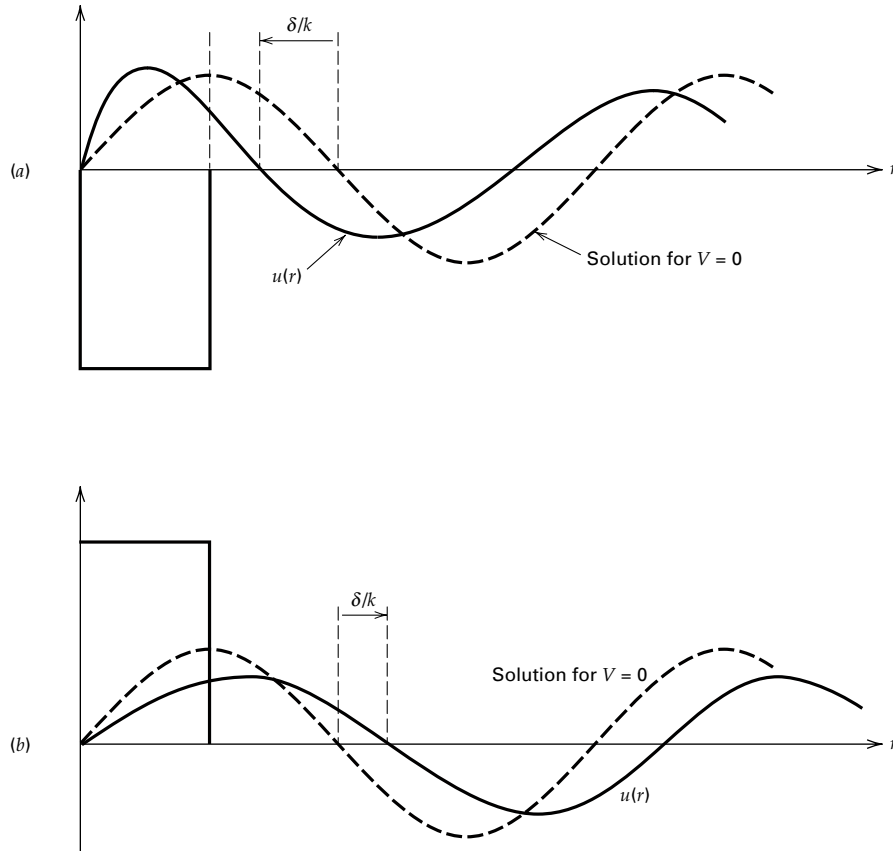


Figure 8B-1 Continuum solution $u(r) = rR_l(r)$ for $l = 0$ (a) attractive potential; (b) repulsive potential.

Comparison with the form in (8B-7), which has the form

$$R_l(r) \rightarrow \frac{\sin(kr - l\pi/2)}{kr} \cos \delta_l(k) + \frac{\cos(kr - l\pi/2)}{kr} \sin \delta_l(k)$$

shows that once we know C/B we can find the phase shift from

$$\frac{C}{B} = -\tan \delta_l(k) \quad (8B-13)$$

The actual calculation of C/B is very tedious, except when $l = 0$. In that case, using $u_l(r) = rR_l(r)$, we just have to match $A \sin kr$ to $B \sin kr + C \cos kr$ (and the derivatives) at $r = a$. Figure (8B-1) shows the shape of the wave functions for attractive and repulsive potentials for $l = 0$.

We conclude with the proof of a useful relation

$$e^{ikr \cos \theta} = \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos \theta) \quad (8B-14)$$

which is of great importance in scattering theory.

The Plane Wave in Terms of Spherical Harmonics

The solution of the free-particle equation

$$\nabla^2 \psi(\mathbf{r}) + k^2 \psi(\mathbf{r}) = 0 \quad (8B-15)$$

can be written in two ways. One is simply the plane wave solution

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \quad (8B-16)$$

The other way is to write it as a linear superposition of the partial wave solutions—that is,

$$\sum \sum A_{lm} j_l(kr) Y_{lm}(\theta, \phi) \quad (8B-17)$$

We may therefore find A_{lm} such that $\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}$ in (8B-17). Note that the spherical angles (θ, ϕ) are the coordinates of the vector $\mathbf{r}$ relative to some arbitrarily chosen z -axis. If we define the z -axis by the direction of $\mathbf{k}$ (until now an arbitrary direction), then

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{ikr \cos \theta} \quad (8B-18)$$

Thus the left side of (8B-18) has no azimuthal angle, ϕ , dependence, and thus on the right side only terms with $m = 0$ can appear; hence, making use of the fact that

$$Y_{l0}(\theta, \phi) = \left(\frac{2l+1}{4\pi} \right)^{1/2} P_l(\cos \theta) \quad (8B-19)$$

where the $P_l(\cos \theta)$ are the Legendre polynomials, we get the relation

$$e^{ikr \cos \theta} = \sum_{l=0}^{\infty} \left(\frac{2l+1}{4\pi} \right)^{1/2} A_l j_l(kr) P_l(\cos \theta) \quad (8B-20)$$

We may use the relation

$$\frac{1}{2} \int_{-1}^1 d(\cos \theta) P_l(\cos \theta) P_{l'}(\cos \theta) = \frac{\delta_{ll'}}{2l+1} \quad (8B-21)$$

which is a direct consequence of the orthonormality relation for the Y_{lm} and (8B-19) to obtain

$$A_l j_l(kr) = \frac{1}{2} [4\pi(2l+1)]^{1/2} \int_{-1}^1 dz P_l(z) e^{ikrz} \quad (8B-22)$$

Compare the two sides of the equation as $kr \rightarrow 0$. The first term on the left-hand side is

$$A_l \frac{(kr)^l}{1, 3, 5, \dots, (2l+1)}$$

and the corresponding power of $(kr)^l$ on the right-hand side has

$$\frac{1}{2} [4\pi(2l+1)]^{1/2} (ikr)^l \int_{-1}^1 dz P_l(z) z^l / l!$$

The integral can be evaluated by noting that $P_l(z)$ is an l th-degree polynomial in z . The coefficient of the leading power, z^l , can be easily obtained from eq. (7-47) as the power of z^l in

$$(-1)^l \frac{1}{2^l l!} \left(\frac{d}{dz} \right)^l (1 - z^2)^l = \frac{2l(2l-1)(2l-2) \cdots (l+1)}{2^l l!} z^l + 0(z^{l-1})$$

W-42 Supplement 8-B The Square Well, Continuum Solutions

We can rewrite this in the form

$$z^l = \frac{2^l l!}{2l(2l-1)(2l-2) \cdots (l+1)} P_l(z) + \text{terms involving } P_{l+1}(z) \text{ and higher}$$

With the help of (8B-21) we finally get

$$A_l \frac{(kr)^l}{1, 3, 5, \dots, (2l+1)} = \frac{1}{2} [4\pi(2l+1)]^{1/2} (ikr)^l \frac{1}{l!} \frac{2^l l!}{2l(2l-1)(2l-2) \cdots (l+1)} \frac{2}{2l+1} \quad (8B-23)$$

What results is the expansion

$$e^{ikr \cos \theta} = \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos \theta)$$

which we will find exceedingly useful in discussions of collision theory.

Chapter 9

Matrix Representation of Operators

The original discovery of quantum mechanics is due to W. Heisenberg. He associated physical quantities like x and p with square *arrays* of numbers for which he proposed multiplication rules, so as to be able to talk about quantities like x^2 and the like. Max Born immediately recognized these arrays as *matrices*, and the formulation of the fundamentals of quantum mechanics was carried by M. Born and P. Jordan with Heisenberg. The discovery of the Schrödinger equation came soon after, followed by the more abstract formulation in terms of operators and states by Paul Dirac. The formulation in terms of matrices is very useful in that it allows us to deal with quantities that have no classical counterpart. This applies particularly to a nonclassical property of particles, their *intrinsic spin*, which will be discussed in the next chapter.

9-1 MATRICES IN QUANTUM MECHANICS

In Chapter 8 we solved the eigenvalue problem posed by the equations

$$\begin{aligned} \mathbf{L}^2|l, m\rangle &= \hbar^2 l(l+1)|l, m\rangle \\ L_z|l, m\rangle &= \hbar m|l, m\rangle \end{aligned} \quad (9-1)$$

by operator methods. We also found that for the harmonic oscillator we could do the same, and even found the complete set of eigenstates in the form

$$|n\rangle = \sqrt{\frac{1}{n!}} (A^+)^n |0\rangle \quad (9-2)$$

for which

$$H|n\rangle = \hbar\omega(n + \frac{1}{2})|n\rangle \quad (9-3)$$

We could also calculate the action of the raising and lowering operators on $|n\rangle$,

$$\begin{aligned} A^+|n\rangle &= \sqrt{n+1}|n+1\rangle \\ A|n\rangle &= \sqrt{n}|n-1\rangle \end{aligned} \quad (9-4)$$

We also showed that

$$\langle m|n\rangle = \delta_{mn} \quad (9-5)$$

a statement that can be made about the eigenstates of *any* hermitian operator.

The expansion theorem was seen to be equivalent to what we called the *completeness* relation

$$\sum_{n=0}^{\infty} |n\rangle\langle n| = \mathbf{1} \quad (9-6)$$

where $\mathbf{1}$ is the unit operator.

We start on the form of matrix representations of operators by practicing on the harmonic oscillator problem. If we take the scalar products of (9-3) and (9-4) with $\langle m|$, we get

$$\begin{aligned} \langle m|H|n\rangle &= \hbar\omega\left(n + \frac{1}{2}\right)\delta_{mn} \\ \langle m|A^+|n\rangle &= \sqrt{n+1}\delta_{m,n+1} \\ \langle m|A|n\rangle &= \sqrt{n}\delta_{m,n-1} \end{aligned} \quad (9-7)$$

These quantities may be arranged in arrays called *matrices*. The conventional notation for a matrix M_{ij} has the first index labeling the row and the second labeling the column of the array. Thus if we write the scalar product $\langle m|H|n\rangle$ as H_{mn} , we find that

$$H = \hbar\omega \begin{pmatrix} 1/2 & 0 & 0 & 0 & \cdots \\ 0 & 3/2 & 0 & 0 & \cdots \\ 0 & 0 & 5/2 & 0 & \cdots \\ 0 & 0 & 0 & 7/2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (9-8)$$

We may also construct matrices for the operator A^+ and A .

$$A^+ = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \sqrt{1} & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (9-9)$$

and

$$A = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & \cdots \\ 0 & 0 & \sqrt{0} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (9-10)$$

Without restricting ourselves to the harmonic oscillator problem, we shall call the array $\langle m|F|n\rangle$, where F is any operator and the $|n\rangle$ any complete set of states, a matrix representation of F in the basis provided by the complete set of states $|n\rangle$. The appellation needs some justification. The product of two matrices, for example, is given by

$$(FG)_{ij} = \sum_k (F)_{ik} (G)_{kj} \equiv (F)_{ik} (G)_{kj} \quad (9-11)$$

[In the last step we have introduced the convention that if a subscript appears twice, it is to be summed over].

We need to verify that this relation holds for the *matrix representations* of the operators F and G . To do this, let us consider some complete set $|n\rangle$. The state $G|j\rangle$ may, with the help of the completeness relation, be written in the form

$$G|j\rangle = \sum_n C_n |n\rangle$$

with

$$C_n = \langle n | G | j \rangle$$

so that

$$G|j\rangle = \sum_n |n\rangle \langle n | G | j \rangle \quad (9-12)$$

Hence

$$\langle i | FG | j \rangle = \sum_n \langle i | F | n \rangle \langle n | G | j \rangle \quad (9-13)$$

This is the same as (9-11), provided we write

$$\langle i | F | n \rangle \equiv F_{in} \quad (9-14)$$

This identification gives rise to the name given to the left-hand side of the above equation. It is called the *matrix element* of the operator F . Further justification for the matrix connection comes from the relation

$$\langle m | F | n \rangle^* = \langle F(n) | m \rangle = \langle n | F^+ | m \rangle \quad (9-15)$$

which shows that if the operator F is represented by a matrix, then the hermitian conjugate operator F^+ will be represented by the hermitian conjugate matrix, since the latter is defined by

$$(F^+)_{nm} = F_{mn}^* \quad (9-16)$$

If the complete set of kets are the eigenkets of some operator A , then the matrix representation of that operator is

$$\langle a_n | A | a_m \rangle = a_m \langle a_n | a_m \rangle = a_m \delta_{mn} \quad (9-17)$$

as is the case in (9-8). Finding the eigenvalues of a hermitian operator is tantamount to *diagonalizing it*.

How do we diagonalize a matrix? Let us consider the operator A , whose eigenkets are $|a_n\rangle$. Suppose we know the *matrix elements* of A in a basis $|u_n\rangle$; that is, we know each entry in the square matrix

$$A_{mn} = \langle u_m | A | u_n \rangle \quad (9-18)$$

We know from the completeness of the basis $|u_n\rangle$ that an arbitrary eigenket $|a_n\rangle$ may be expanded in terms of the $|u_n\rangle$. We write

$$|a_n\rangle = \sum_k |u_k\rangle \langle u_k | a_n \rangle \quad (9-19)$$

and similarly

$$\langle a_n | = \sum_l \langle a_n | u_l \rangle \langle u_l | \quad (9-20)$$

It follows that

$$\langle a_m | A | a_n \rangle = a_m \delta_{mn} = \sum_l \langle a_m | u_l \rangle \langle u_l | A | u_k \rangle \langle u_k | a_n \rangle \quad (9-21)$$

Let us now define the matrix U by the matrix elements

$$U_{kn} \equiv \langle u_k | a_n \rangle \quad (9-22)$$

Then

$$\langle a_m | u_l \rangle = \langle u_l | a_m \rangle^* = U_{lm}^* = (U^+)_{ml} \quad (9-23)$$

Thus, in matrix form the equation reads

$$a_m \delta_{mn} = \sum_{k,l} (U^+)_{ml} A_{lk} U_{kn} \quad (9-24)$$

The matrix U has the property that

$$\sum_k \langle a_m | u_k \rangle \langle u_k | a_n \rangle = (U^+ U)_{mn} = \delta_{mn} \quad (9-25)$$

that is,

$$U^+ U = 1 \quad (9-26)$$

Such a matrix is called a *unitary* matrix. Finding the matrix U for finite-dimensional matrices A is an exercise in solving algebraic equations. Finding it for infinite-dimensional A 's is equivalent to solving the Schrödinger equation.

We shall find it useful to introduce the *trace* of a matrix. This is defined by

$$\text{Tr } M = \sum_k M_{kk} \quad (9-27)$$

The trace of a product of two matrices has the interesting property that

$$\begin{aligned} \text{Tr } MN &= \sum_k (MN)_{kk} = \sum_{kl} M_{kl} N_{lk} = \sum_{kl} N_{lk} M_{kl} \\ &= \text{Tr } NM \end{aligned} \quad (9-28)$$

Note that we were allowed to order the matrix elements differently because each one of the M_{kl} and N_{kl} are plain (complex) numbers. It is easy to show that

$$\text{Tr } ABC = \text{Tr } BCA = \text{Tr } CAB \quad (9-29)$$

and similarly for products of larger number of matrices. A consequence of this is that

$$\text{Tr } [A, B] = 0 \quad (9-30)$$

unless the series in (9-27) diverges, in which case no useful statement can be made. Thus, for finite-dimensional matrices, such as the matrix representation of angular momentum,

$$[L_x, L_y] = i\hbar L_z \quad (\text{and cycl.}),$$

implies that the traces of the matrices for $\mathbf{L}$ must vanish. We can turn such an argument around and argue that for operators such as x_{op} and p_{op} , the relation

$$[p_{\text{op}}, x_{\text{op}}] = \frac{\hbar}{i} \mathbf{1}$$

implies that these operators *must* be represented by infinite-dimensional matrices. For example in the harmonic oscillator basis, we have seen that A and A^+ are represented by infinite-dimensional matrices, and therefore so are the x and p operators.

9-2 MATRIX REPRESENTATIONS OF ANGULAR MOMENTUM OPERATORS

It is often convenient and, as we shall see in the next chapter, sometimes *essential*, to represent operators as matrices. Let us do this for the angular momentum operators. Consider the matrix elements of L_z . As a consequence of

$$[\mathbf{L}^2, L_z] = 0$$

we have

$$\begin{aligned} 0 &= \langle l', m' | [\mathbf{L}^2, L_z] | l, m \rangle \\ &= \langle \mathbf{L}^2(l', m') | L_z | l, m \rangle - \langle l', m' | L_z \mathbf{L}^2 | l, m \rangle \\ &= \hbar^2[l'(l' + 1) - l(l + 1)] \langle l', m' | L_z | l, m \rangle \end{aligned} \quad (9-31)$$

From this we conclude that if $l' \neq l$ then $\langle l', m' | L_z | l, m \rangle$ vanishes. Thus L_z (and all other operators that commute with $\mathbf{L}^2$) only have matrix elements between states that have the same total angular momentum quantum numbers. If we stay with fixed l , then we use the second of the relations in (9-1) to get

$$\langle l, m' | L_z | l, m \rangle = \hbar m \delta_{m'm} \quad (9-32)$$

Furthermore, (9-16) and (9-18) imply that

$$\langle l, m' | L_{\pm} | l, m \rangle = \hbar \sqrt{l(l + 1) - m(m \pm 1)} \delta_{m', m \pm 1} \quad (9-33)$$

For the angular momentum $l = 1$ operators, this leads to the matrix representations

$$L_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (9-34)$$

We also get

$$L_+ = \hbar \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \quad (9-35)$$

and

$$L_- = \hbar \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \quad (9-36)$$

From these we get

$$L_x = \frac{1}{2}(L_+ + L_-) = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (9-37)$$

and

$$L_y = -\frac{i}{2}(L_+ - L_-) = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad (9-38)$$

The rows and columns are labeled with $m = 1, 0, -1$ in order from left to right and top to bottom. The matrices satisfy the commutation relations: for example,

$$\begin{aligned} [L_x, L_y] &= i \frac{\hbar^2}{2} \left[\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right] \\ &= i \frac{\hbar^2}{2} \left[\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix} - \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \right] = i \frac{\hbar^2}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} \\ &= i\hbar L_z \end{aligned}$$

9-3 GENERAL RELATIONS IN MATRIX MECHANICS

Consider, for example, a relation between state vectors that reads as follows:

$$|\psi\rangle = A|\phi\rangle \quad (9-39)$$

If we take the scalar product of this with any member of a complete set $|i\rangle$, we have

$$\langle i|\psi\rangle = \langle i|A|\phi\rangle \quad (9-40)$$

Furthermore, the insertion of the unit operator in the form (6-16) between A and $|\phi\rangle$ yields

$$\langle i|\psi\rangle = \sum_n \langle i|A|n\rangle \langle n|\phi\rangle \quad (9-41)$$

If we write $\langle n|\phi\rangle$ as a column vector α_n , so that

$$\langle n|\phi\rangle \rightarrow \begin{pmatrix} \langle 1|\phi\rangle \\ \langle 2|\phi\rangle \\ \langle 3|\phi\rangle \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \end{pmatrix} \quad (9-42)$$

and similarly

$$\langle n|\psi\rangle \rightarrow \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \langle 3|\psi\rangle \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \end{pmatrix} \quad (9-43)$$

then the matrix representation of (9-39) is

$$\beta_i = \sum_n A_{in} \alpha_n \quad (9-44)$$

Thus matrices represent operators, and column vectors represent kets. The scalar product $\langle\phi|n\rangle = \langle n|\phi\rangle^*$ is conventionally written in the form of a row,

$$\langle\phi|n\rangle \rightarrow (\alpha_1^*, \alpha_2^*, \alpha_3^*, \dots) \quad (9-45)$$

so that the scalar product $\langle\phi|\psi\rangle$, for example, can be written as

$$\langle\phi|\psi\rangle = \sum_n \langle\phi|n\rangle \langle n|\psi\rangle = \sum_n \alpha_n^* \beta_n \quad (9-46)$$

Let us now calculate an expectation value of some operator Q in a state $|\psi\rangle$. The expectation value is $\langle\psi|Q|\psi\rangle$, and this may be translated into a relation involving matrices and column vectors:

$$\langle\psi|Q|\psi\rangle = \sum_{m,n} \langle\psi|n\rangle\langle n|Q|m\rangle\langle m|\psi\rangle = \sum_{m,n} \beta_n^* Q_{nm} \beta_m \quad (9-47)$$

More generally, for the calculation of a more general *matrix element*, we have,

$$\begin{aligned} \langle\phi|Q|\psi\rangle &= \sum_{m,n} \langle\phi|n\rangle\langle n|Q|m\rangle\langle m|\psi\rangle = \sum_{m,n} \alpha_n^* Q_{nm} \beta_m \\ &= (\alpha_1^*, \alpha_2^*, \alpha_3^* \dots) \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} & \cdots \\ Q_{21} & Q_{22} & Q_{23} & \cdots \\ Q_{31} & Q_{32} & Q_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \end{pmatrix} \end{aligned} \quad (9-48)$$

An eigenvalue equation is a special case of (9-39). It reads

$$A|\phi\rangle = a|\phi\rangle \quad (9-49)$$

so that in matrix form it represents a set of algebraic equations

$$\sum_n A_{in} \alpha_n = a \alpha_i \quad (9-50)$$

This is equivalent to

$$\begin{pmatrix} A_{11} - a & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} - a & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} - a & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \end{pmatrix} = 0 \quad (9-51)$$

We use a theorem from the theory of matrices to assert that there will be a nontrivial solution of this equation only if the determinant of the matrix vanishes:

$$\det|A_{in} - a \delta_{in}| = 0 \quad (9-52)$$

This is a good way of finding eigenvalues (and eigenvectors) for operators represented by finite matrices. For infinite matrices, however, this is not so simple: It is equivalent to solving the Schrödinger equation.

All of the properties of operators and their associated kets and bras hold for matrices, column vectors, and the complex conjugate row-vectors. Thus, for example,

1. Hermitian matrices have real eigenvalues, and
2. The vectors corresponding to different eigenvalues are orthogonal, so that if the set $\alpha^{(n)}$ corresponds to the eigenvalue a_n , then

$$\sum_i (\alpha^{(n)})_i^* (\alpha^{(m)})_i = \delta_{nm} \quad (9-53)$$

In writing this we have chosen the eigenvectors for each eigenvalue to be normalized.

We can find the form of the unitary matrix U that diagonalizes the matrix A . We did this before in terms of the bras and kets, and we now do this in terms of the eigenvectors of the matrices. We write the eigenvalue equation explicitly:

$$\sum_j A_{ij} \alpha_j^{(n)} = a_n \alpha_i^{(n)} \quad (9-54)$$

Consider now a matrix U , whose matrix elements are defined as follows:

$$U_{in} = \alpha_i^{(n)} \quad (9-55)$$

It follows that

$$(U^+)_{mi} = (U_{in})^* = (\alpha_i^{(n)})^* \quad (9-56)$$

We may therefore write

$$\begin{aligned} (U^+AU)_{mn} &= \sum_{i,j} (U^+)_{mi} A_{ij} U_{jn} = \sum_{i,j} (\alpha_i^{(m)})^* A_{ij} \alpha_j^{(n)} \\ &= \sum_i (\alpha_i^{(m)})^* a_n \alpha_i^{(n)} = a_n \sum_i (\alpha_i^{(m)})^* \alpha_i^{(n)} \\ &= a_n \delta_{mn} \end{aligned} \quad (9-57)$$

This is just a statement that the unitary matrix U^+ , defined above, diagonalizes A . (Note that we could have multiplied each of the columns in the matrix U by a different phase factor of the form $e^{i\phi_n}$, since this would just multiply the result above by $e^{i(\phi_n - \phi_n)}$, which does not affect the final result. We therefore ignore this possibility).

Note that the eigenvectors of a diagonal matrix are particularly simple. If we write

$$\begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & a_4 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \lambda \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

(taking a 4×4 matrix for simplicity), we see immediately that the eigenvalues λ are a_1, a_2, a_3, a_4 , and the corresponding (normalized) eigenvectors are

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

respectively. Note that these are manifestly orthogonal.

We illustrate the manipulation of matrices by the following example:

EXAMPLE 9-1

Consider the matrix

$$M_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(a) Find the eigenvalues and the normalized eigenvectors of the matrix M_3 ; that is, find the values of α_i for which $(M_3)_{ij} \alpha_j = \lambda \alpha_i$.

(b) Find the unitary matrix U that diagonalizes M_3 .

SOLUTION We begin by writing out the equation $(M_3)_{ij} \alpha_j = \lambda \alpha_i$:

$$\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \lambda \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$$

We get $-i\alpha_2 = \lambda\alpha_1$, $i\alpha_1 = \lambda\alpha_2$, and $\lambda\alpha_3 = 0$. From the third of these, we see that if $\lambda \neq 0$, then $\alpha_3 = 0$. Since the product of the first two relations implies $(\lambda^2 - 1)\alpha_1\alpha_2 = 0$, we see that unless $\lambda = +1, -1$, or 0 , all the α 's must vanish. So, for $\lambda = 0$, we get $\alpha_1 = \alpha_2 = 0$, and $\alpha_3 = 1$ since the normalization condition implies that $|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2 = 1$. For $\lambda = 1$, we get $\alpha_1 = u$, $\alpha_2 = iu$, and $\alpha_3 = 0$. Normalization implies that $|u|^2 + |iu|^2 = 2|u|^2 = 1$ so we may choose $u = 1/\sqrt{2}$. For $\lambda = -1$, we do the same thing, or we recognize that eigenvectors corresponding to different eigenvalues are orthogonal. We thus write $\alpha_1 = 1/\sqrt{2}$, $\alpha_2 = -i/\sqrt{2}$, $\alpha_3 = 0$. So for the eigenstates corresponding to the three eigenvalues we get

$$\lambda = 1, \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \\ 0 \end{pmatrix}; \quad \lambda = 0, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \quad \lambda = -1, \begin{pmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \\ 0 \end{pmatrix};$$

We may use this to find U according to the rule given in (9-55):

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

We check:

$$\begin{aligned} U^\dagger M_3 U &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ i & 0 & -i \\ 0 & \sqrt{2} & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \frac{L_z}{\hbar} \end{aligned}$$

Equivalently, M_3 may be seen to be a unitary transform of L_z .

If we can solve the eigenvalue problem for some hermitian operator H , which is easy (and useful) for 2×2 and 3×3 matrices, then it is straightforward to find an expression for any function of H . In general, if

$$U^\dagger H U = D \quad (9-58)$$

where D is a diagonal matrix, then

$$\begin{aligned} \sum_n a_n H^n &= \sum_n a_n (U D U^\dagger)^n = \sum_n a_n (U D U^\dagger)(U D U^\dagger)(U D U^\dagger) \dots \\ &= U \sum_n a_n D^n U^\dagger = U \begin{pmatrix} \sum_n a_n h_1^n & 0 & 0 & \dots \\ 0 & \sum_n a_n h_2^n & 0 & \dots \\ 0 & 0 & \sum_n a_n h_3^n & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} U^\dagger \quad (9-59) \end{aligned}$$

Here the h_k are the eigenvalues of H and therefore the diagonal elements of the matrix D . A simple diagonalization of H yields the eigenvalues and the eigenvectors, which are used to construct the U matrix. We illustrate this in the next example.

EXAMPLE 9-2

Use the procedure of the previous example to calculate the matrix representation of

$$e^{i\alpha M_3} = \sum_{k=0}^{\infty} \frac{(i\alpha)^k}{k!} (M_3)^k$$

SOLUTION We have

$$e^{i\hbar\alpha M_3} = \sum_{k=0}^{\infty} \frac{(i\hbar\alpha)^k}{k!} (M_3)^k$$

We know that

$$U^\dagger M_3 U = L_z / \hbar$$

It follows that

$$\begin{aligned} (L_z / \hbar)^k &= (U^\dagger M_3 U)^k = U^\dagger M_3 U U^\dagger M_3 U \dots U^\dagger M_3 U \\ &= U^\dagger (M_3)^k U \end{aligned}$$

Equivalently,

$$(M_3)^k = U (L_z / \hbar)^k U^\dagger$$

Now for k even,

$$(L_z / \hbar)^k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

While for k odd,

$$(L_z / \hbar)^k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = (L_z / \hbar)$$

We use this to evaluate the sum

$$\sum_{k=0}^{\infty} \frac{(i\alpha M_3)^k}{k!} = U \sum_{k=0}^{\infty} \frac{(i\alpha L_z / \hbar)^k}{k!} U^\dagger$$

The calculation of the sum is straightforward:

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{(i\alpha L_z / \hbar)^k}{k!} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{2!} \begin{pmatrix} \alpha^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha^2 \end{pmatrix} + \frac{1}{4!} \begin{pmatrix} \alpha^4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha^4 \end{pmatrix} - \dots \\ &\quad + i \left[\begin{pmatrix} \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\alpha \end{pmatrix} - \frac{1}{3!} \begin{pmatrix} \alpha^3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\alpha^3 \end{pmatrix} + \frac{1}{5!} \begin{pmatrix} \alpha^5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\alpha^5 \end{pmatrix} + \dots \right] \\ &= \begin{pmatrix} \cos \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \cos \alpha \end{pmatrix} + i \begin{pmatrix} \sin \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\sin \alpha \end{pmatrix} \\ &= \begin{pmatrix} e^{i\alpha} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-i\alpha} \end{pmatrix} \end{aligned}$$

What remains is the calculation of

$$U \begin{pmatrix} e^{i\alpha} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-i\alpha} \end{pmatrix} U^\dagger$$

an exercise left to the reader.

PROBLEMS

1. Given the matrix representation for A^+ in (9-9), what are the matrix forms for $(A^+)^n$ for $n = 2, 3, 4$? (Keep only the left-hand 4×4 corner of the infinite array.)
2. Obtain a matrix representation for the operators x and x^2 , using (9-9) and (9-10). [See eq. (6-36).]
3. Obtain a matrix representation for the operators p and p^2 , using (9-9) and (9-10). [See eq. (6-36).]
4. Suppose the ground state $|0\rangle$ that is acted on by the matrices $(A^+)^n$ is represented by the column vector

$$u_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}$$

Use (9-2) and (9-9) to obtain the column vector u_1 first excited state, $|1\rangle$.

5. Use the results of Problem 1 to obtain the column vectors for $|2\rangle$, $|3\rangle$, and $|4\rangle$. Do you see a pattern there?
6. Given the state vector represented by the column

$$\psi = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \\ \vdots \end{pmatrix}$$

Calculate the following quantities, as defined by the harmonic oscillator:

- (a) $\langle H \rangle$, in that state
 - (b) $\langle x^2 \rangle$, $\langle x \rangle$, $\langle p^2 \rangle$, $\langle p \rangle$, in that state.
 - (c) Use your results in (b) to calculate $(\Delta x)^2$ and $(\Delta p)^2$.
7. You are given the hermitian matrix

$$A = \begin{pmatrix} -3 & \sqrt{\frac{19}{4}} e^{i\pi/3} \\ \sqrt{\frac{19}{4}} e^{-i\pi/3} & 6 \end{pmatrix}$$

- (a) Find the eigenvalues.
 - (b) Find the eigenvectors.
 - (c) Find the matrix U that diagonalizes A .
8. Use the results of Problem 7 to calculate e^{-A} .
 9. Consider the matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Calculate the eigenvalues and the eigenvectors of this matrix.

10. Consider an arbitrary hermitian matrix A . Show that

$$\text{Tr } A = \sum_{i=1}^N A_{ii}$$

is just the sum of the eigenvalues of A . [Hint: Use eq. (9-29).]

11. Consider an $N \times N$ matrix, which has a 1 in every entry (a generalization of the matrix in Problem 9). Find its eigenvalues by calculating the square of the matrix and by using the result of Problem 10.
12. In Example 9-1 we found the matrix U , which has the property that $U^\dagger M_3 U = L_z/\hbar$ so that $U L_z U^\dagger = \hbar M_3$.
 - (a) Calculate M_1 and M_2 by applying U and $U^\dagger$ to L_x and L_y , respectively.
 - (b) Show by explicit calculation that the M 's satisfy the angular momentum commutation relations.
13. Consider two hermitian matrices A and B that commute, so that $[A, B] = 0$. Let U be a unitary matrix that diagonalizes A . Show that if the eigenvalues of A are all different then the same matrix U also diagonalizes B . (*Hint:* A simple way to see this is to work with a finite—say, 4×4 —matrix and work things out explicitly).
14. Consider an arbitrary matrix M . Show that if $[M, M^\dagger] = 0$ then M can be diagonalized by a unitary matrix. (*Hint:* Consider the matrices $M + M^\dagger$, $i(M - M^\dagger)$ and use your result in Problem 13.)

Chapter 10

Spin

A proper discussion of atomic and subatomic physics is not possible without taking into account a new degree of freedom for the electron—the *spin* of the electron. In spite of its suggestive name, this property of the electron does not have a classical analog. The picture of an electron rotating about its own axis is wrong. Thus there is no counterpart to the orbital angular momentum operator $\mathbf{r} \times \mathbf{p}$, and the treatment must necessarily be more abstract. It turns out that the operator treatment of angular momentum, which depends only on the commutation relations, also applies to spin. It is only the step of representing the states by wave functions whose variables are real coordinates (such as θ, φ) that cannot be carried out here. Rather, the operators are represented by matrices.

The discovery of spin was rooted in experiment. There were many mysterious patterns that appeared in the pre-quantum analysis of spectral lines. Among them was the appearance of doublets of lines when radiating alkali atoms were placed in a magnetic field. The semiclassical understanding of what was called the *Zeeman effect* after its discoverer did not allow for such splittings, and it took a while before W. Pauli in 1924 (i.e., before quantum mechanics) saw through the fog and proposed that an electron in an atom, besides being described by a set of quantum numbers such as (n, l, m) , had another nonclassical two-valued “index” associated with it. Within a year Samuel Goudsmit and George Uhlenbeck, at the time graduate students in Holland, proposed that the electron had an intrinsic angular momentum that had the value $s = 1/2$. The electron is pointlike, and there is nothing that rotates. The spin of the electron is just an intrinsic property like its mass, and the *spin operators* in analogy with the orbital angular momentum discussed in Chapter 7 are defined by the commutation relations

$$[S_x, S_y] = i\hbar S_z \quad (\text{and cycl.}) \quad (10-1)$$

10-1 EIGENSTATES OF SPIN 1/2

We recall that eqs. (9-32) and (9-33) were derived solely from the commutation relations like those of (10-1). Applying them to the spin case, where $s = 1/2$, we can write

$$S_z = \hbar \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \quad (10-2)$$

and (9-33) gives

$$S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (10-3)$$

We may write this representation as

$$\mathbf{S} = \frac{1}{2}\hbar\boldsymbol{\sigma} \quad (10-4)$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (10-5)$$

are the *Pauli matrices*. They satisfy the commutation relations

$$[\sigma_x, \sigma_y] = 2i\sigma_z \quad (10-6)$$

and so on, as they must, to satisfy (10-1), and they also satisfy

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \equiv 1 \quad (10-7)$$

The Pauli matrices also anticommute:

$$\begin{aligned} \sigma_x \sigma_y &= -\sigma_y \sigma_x \\ \sigma_z \sigma_x &= -\sigma_x \sigma_z \\ \sigma_y \sigma_z &= -\sigma_z \sigma_y \end{aligned} \quad (10-8)$$

These relations are peculiar to the spin 1/2 representations and do not hold for the $l = 1$ matrices, for example.

The eigenstates of S_z will be represented by a two-component column vector, which we call *spinor*. To find these eigenspinors, we solve

$$S_z \begin{pmatrix} u \\ v \end{pmatrix} = \pm \frac{1}{2} \hbar \begin{pmatrix} u \\ v \end{pmatrix} \quad (10-9)$$

that is, using (10-4),

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \pm \begin{pmatrix} u \\ v \end{pmatrix}$$

or

$$\begin{pmatrix} u \\ -v \end{pmatrix} = \pm \begin{pmatrix} u \\ v \end{pmatrix}$$

The plus eigensolution has $v = 0$, and the minus eigensolution has $u = 0$. We thus write

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (10-10)$$

for the eigenspinors corresponding to spin up [$S_z = +(1/2)\hbar$] and spin down [$S_z = -(1/2)\hbar$], respectively.

An arbitrary spinor can be expanded in this complete set

$$\begin{pmatrix} \alpha_+ \\ \alpha_- \end{pmatrix} = \alpha_+ \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_- \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (10-11)$$

and the expansion postulate yields the interpretation that $|\alpha_+|^2$ and $|\alpha_-|^2$, when properly normalized, so that

$$|\alpha_+|^2 + |\alpha_-|^2 = 1 \quad (10-12)$$

yield the probabilities that a measurement of S_z on the state $\begin{pmatrix} \alpha_+ \\ \alpha_- \end{pmatrix}$ yields $+(1/2)\hbar$ and $-(1/2)\hbar$, respectively.

It is not necessary to keep S_z diagonal. If we look for eigenstates of the operator $S_x \cos \phi + S_y \sin \phi$, we must solve

$$(S_x \cos \phi + S_y \sin \phi) \begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{2} \hbar \lambda \begin{pmatrix} u \\ v \end{pmatrix} \quad (10-13)$$

that is,

$$\begin{pmatrix} 0 & \cos \phi - i \sin \phi \\ \cos \phi + i \sin \phi & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \end{pmatrix}$$

This implies that

$$\begin{aligned} v e^{-i\phi} &= \lambda u \\ u e^{i\phi} &= \lambda v \end{aligned} \quad (10-14)$$

Taking the products of the left- and right-hand sides in the two equations we find that

$$uv(\lambda^2 - 1) = 0 \quad (10-15)$$

Hence

$$\lambda = \pm 1 \quad (10-16)$$

The eigenvector corresponding to $\lambda = 1$ satisfies

$$v = e^{i\phi} u$$

so that the normalized form is

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\phi} \end{pmatrix}$$

We take advantage of the fact that we can multiply a state vector by an arbitrary phase factor, which we choose to be $e^{-i\phi/2}$. This yields

$$u_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\phi/2} \\ e^{i\phi/2} \end{pmatrix} \quad (10-17)$$

and similarly, the eigenstate corresponding to $\lambda = -1$ can be written in the form

$$u_- = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\phi/2} \\ -e^{i\phi/2} \end{pmatrix} \quad (10-18)$$

which is easily seen to be orthogonal to u_+ :

$$\begin{aligned} u_+^* u_- &= \frac{1}{2} (e^{i\phi/2}, e^{-i\phi/2}) \begin{pmatrix} e^{-i\phi/2} \\ -e^{i\phi/2} \end{pmatrix} \\ &= 0 \end{aligned} \quad (10-19)$$

Given an arbitrary state α , the expectation value of $\mathbf{S}$ can be calculated. We have

$$\langle \alpha | \mathbf{S} | \alpha \rangle = \sum_i \sum_j \langle \alpha | i \rangle \langle i | \mathbf{S} | j \rangle \langle j | \alpha \rangle$$

or, equivalently,

$$(\alpha_+^*, \alpha_-^*) \mathbf{S} \begin{pmatrix} \alpha_+ \\ \alpha_- \end{pmatrix}$$

Thus

$$\begin{aligned}
 \langle S_x \rangle &= (\alpha_+^*, \alpha_-^*) \frac{1}{2} \hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_+ \\ \alpha_- \end{pmatrix} \\
 &= \frac{1}{2} \hbar (\alpha_+^*, \alpha_-^*) \begin{pmatrix} \alpha_- \\ \alpha_+ \end{pmatrix} = \frac{1}{2} \hbar (\alpha_+^* \alpha_- + \alpha_-^* \alpha_+) \\
 \langle S_y \rangle &= \frac{1}{2} \hbar (\alpha_+^*, \alpha_-^*) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \alpha_+ \\ \alpha_- \end{pmatrix} \\
 &= \frac{1}{2} \hbar (\alpha_+^*, \alpha_-^*) \begin{pmatrix} -i\alpha_- \\ i\alpha_+ \end{pmatrix} = -\frac{i\hbar}{2} (\alpha_+^* \alpha_- - \alpha_-^* \alpha_+) \\
 \langle S_z \rangle &= \frac{1}{2} \hbar (\alpha_+^*, \alpha_-^*) \begin{pmatrix} \alpha_+ \\ -\alpha_- \end{pmatrix} = \frac{1}{2} \hbar (|\alpha_+|^2 - |\alpha_-|^2)
 \end{aligned} \tag{10-20}$$

Note that all of these are real, as expected for hermitian operators.

EXAMPLE 10-1

A measurement of S_x for a spin 1/2 system leads to the eigenvalue $+\hbar/2$. Subsequently, a measurement of $S_x \cos \phi + S_y \sin \phi$ is carried out. What is the probability that the result is $+\hbar/2$?

SOLUTION After the initial measurement, the system must be in the eigenstate of S_x with eigenvalue $+\hbar/2$. This means that

$$\frac{\hbar}{2} \sigma_x \begin{pmatrix} u \\ v \end{pmatrix} = + \frac{\hbar}{2} \begin{pmatrix} u \\ v \end{pmatrix}$$

that is,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$$

Since $|u|^2 + |v|^2 = 1$ results from normalization, it is convenient to choose the numbers u and v to be real. Thus

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

This state may be expanded in eigenstates of the operator $S_x \cos \phi + S_y \sin \phi$, which were obtained before. They are given in (10-17) and (10-18). Writing

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \alpha_+ \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\phi/2} \\ e^{i\phi/2} \end{pmatrix} + \alpha_- \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\phi/2} \\ -e^{i\phi/2} \end{pmatrix}$$

we get

$$\alpha_+ = \frac{1}{2} (e^{i\phi/2} \ e^{-i\phi/2}) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \cos \frac{\phi}{2}$$

The result is that $P_+ = |\alpha_+|^2 = \cos^2 \phi/2$.

10-2 THE INTRINSIC MAGNETIC MOMENT OF SPIN 1/2 PARTICLES

The existence of spin has immediate consequences for physical systems. We shall see later that the spin of the electron alters the Hamiltonian for the hydrogen atom (and by extension, for other atoms). The reason is that the electron has an intrinsic magnetic dipole moment by virtue of its spin. Under certain circumstances it is possible to treat the electron spin as the only degree of freedom that an electron possesses: This happens when an electron is localized inside a crystal lattice. When this is the case, the coupling of the magnetic moment to an externally imposed magnetic field has consequences that will be explored in this and the next section.

The magnetic moment is¹

$$\mathbf{M} = -\frac{eg}{2m_e} \mathbf{S} \quad (10-21)$$

Where g , the gyromagnetic ratio, is very close to 2.² For such a localized electron, the Hamiltonian in the presence of a magnetic field $\mathbf{B}$ is just the potential energy

$$H = -\mathbf{M} \cdot \mathbf{B} = \frac{eg\hbar}{4m_e} \boldsymbol{\sigma} \cdot \mathbf{B} \quad (10-22)$$

The Schrödinger equation for the state $\psi(t) = \begin{pmatrix} \alpha_+(t) \\ \alpha_-(t) \end{pmatrix}$ is

$$i\hbar \frac{d}{dt} \psi(t) = \frac{eg\hbar}{4m_e} \boldsymbol{\sigma} \cdot \mathbf{B} \psi(t) \quad (10-23)$$

If $\mathbf{B}$ is taken to define the z -axis, and if we write $\psi(t) = e^{-i\omega t} \begin{pmatrix} \alpha_+ \\ \alpha_- \end{pmatrix}$, then the equation becomes

$$\hbar\omega \begin{pmatrix} \alpha_+ \\ \alpha_- \end{pmatrix} = \frac{eg\hbar B}{4m_e} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha_+ \\ \alpha_- \end{pmatrix}$$

One eigenvalue is $\omega_0 = egB/4m_e$ and the corresponding solution is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. The other eigenvalue is $\omega_0 = -egB/4m_e$ and the eigenvector is $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. If the initial state is

$$\psi(0) = \begin{pmatrix} a \\ b \end{pmatrix} \quad (10-24)$$

then the state at a later time will be

$$\psi(t) = \begin{pmatrix} ae^{-i\omega_0 t} \\ be^{i\omega_0 t} \end{pmatrix} \quad (10-25)$$

¹A classical charge $-e$ moving in a circle with angular momentum $\mathbf{L}$ will form a current loop whose magnetic moment is $\mathbf{M} = -e\mathbf{L}/2m$. Since spin has no “classical limit,” the justification for (10-21) has to be found elsewhere. The relativistic Dirac equation yields this result, as well as the value $g = 2$.

²The value of $g = 2(1.0011596 \dots)$ is one of the most accurately measured (and calculated) numbers in all of physics. The theoretical and experimental values agree to the last decimal place, with the only theoretical uncertainties coming from the still poorly understood short-distance properties of the fundamental particles.

Suppose that at time $t = 0$ the spin is in an eigenstate of S_x with eigenvalue $\hbar/2$, so that it “points in the x -direction.” This means that

$$\frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} a \\ b \end{pmatrix}$$

that is,

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (10-26)$$

At a later time,

$$\psi(t) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\omega_0 t} \\ e^{i\omega_0 t} \end{pmatrix} \quad (10-27)$$

At the time t ,

$$\langle S_x \rangle = \frac{\hbar}{2} \frac{1}{\sqrt{2}} (e^{i\omega_0 t} \ e^{-i\omega_0 t}) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\omega_0 t} \\ e^{i\omega_0 t} \end{pmatrix} = \frac{\hbar}{2} \cos 2\omega_0 t \quad (10-28)$$

Similarly,

$$\langle S_y \rangle = \frac{\hbar}{2} \frac{1}{\sqrt{2}} (e^{i\omega_0 t} \ e^{-i\omega_0 t}) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\omega_0 t} \\ e^{i\omega_0 t} \end{pmatrix} = \frac{\hbar}{2} \sin 2\omega_0 t \quad (10-29)$$

Thus the spin precesses about the direction of B with a frequency

$$2\omega_0 = \frac{egB}{2m_e} \equiv g\omega_c \quad (10-30)$$

$\omega_c = eB/2m_e$ is called the *cyclotron frequency*. The same frequency of precession occurs if the spin makes an angle θ with the z -axis, rather than lying in the x -plane. For a magnetic field of 1T, the magnitude of $\omega_c = 0.9 \times 10^{11}$ radians/s.

It is worth commenting that the “motion” of the spin can easily be interpreted when written in the Heisenberg picture. There we have

$$\frac{d\mathbf{S}}{dt} = \frac{i}{\hbar} [H, \mathbf{S}(t)] \quad (10-31)$$

With

$$H = \frac{eg}{2m_e} \mathbf{S} \cdot \mathbf{B} \equiv \gamma \mathbf{S} \cdot \mathbf{B} \quad (10-32)$$

we can calculate

$$\begin{aligned} \frac{dS_1(t)}{dt} &= \frac{i\gamma}{\hbar} [B_1[S_1, S_1] + B_2[S_2, S_1] + B_3[S_3, S_1]] \\ &= \frac{i\gamma}{\hbar} [-i\hbar B_2 S_3 + i\hbar B_3 S_2] = \gamma(\mathbf{B} \times \mathbf{S})_1 \end{aligned}$$

Similar calculation for the other components yields

$$\frac{d\mathbf{S}(t)}{dt} = \gamma(\mathbf{S} \times \mathbf{B}) \quad (10-33)$$

This, however, is exactly the equation for the precession of a classical angular momentum vector (associated with a magnetic dipole) subject to a torque due to a magnetic field. In this form we easily see that the precession is not restricted to the case of spin 1/2. Given that the magnetic moment of a system with angular momentum $\mathbf{J}$ is $\gamma\mathbf{J}$, the same equation holds.

10-3 PARAMAGNETIC RESONANCE

In a solid the gyromagnetic factor g of an electron is affected by the nature of the forces acting in the solid. A knowledge of g provides very useful constraints on what these forces could be, and it is therefore important to be able to measure g . This can be done by the *paramagnetic resonance method*. The principle of the method is the following: We have a magnetic field pointing in the z -direction, and the electron spin precesses about that direction. How fast does it do so? If we could introduce a magnetic field that is perpendicular to the z -axis and rotates with the spin, then the field would “see” an electron spin at rest. The component of the electron spin that points in the x - y plane would preferentially align in a direction opposite to the magnetic field to reach the minimum energy state. For those electrons not already aligned in the minimum energy direction, a transition to the lowest energy will take place, and in the process, energy—in the form of radiation—is given up. This can be detected and measured.

It is not practical to have a magnetic field rotating with a frequency of the order of 10^{11} radians per second. If, however, we have a magnetic field that points in the x -direction, say, and oscillates with a frequency ω , it may be viewed as a superposition of a field rotating in the x - y plane clockwise with frequency ω , and a field rotating counterclockwise with the same frequency, with the phase arranged so that the net effect is in the x -direction. (This is analogous to obtaining a linear polarization out of the sum of two circular polarizations.) Only one of the components will travel in the same direction as the precessing spin. The other component will move in a direction opposite to the spin precession, and its effect on the electron spin averages out to zero.

Consider an electron whose only degrees of freedom are the spin states, under the influence of a large magnetic field B_0 pointing in the z -direction and constant in time, and a small oscillating field $B_1 \cos \omega t$ pointing in the x -direction. The Schrödinger equation now reads

$$i\hbar \frac{d}{dt} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \frac{eg\hbar}{4m_e} \begin{pmatrix} B_0 & B_1 \cos \omega t \\ B_1 \cos \omega t & -B_0 \end{pmatrix} \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \quad (10-34)$$

If $B_1 = 0$,

$$\begin{aligned} i \frac{da(t)}{dt} &= \frac{egB_0}{4m_e} a(t) \equiv \omega_0 a(t) \\ i \frac{db(t)}{dt} &= -\omega_0 b(t) \end{aligned}$$

so that $a(t) = a(0)e^{-i\omega_0 t}$ and $b(t) = b(0)e^{i\omega_0 t}$. This suggests that if $B_1 \neq 0$ then we introduce $A(t)$ and $B(t)$ as follows:

$$\begin{aligned} A(t) &= a(t)e^{i\omega_0 t} \\ B(t) &= b(t)e^{-i\omega_0 t} \end{aligned} \quad (10-35)$$

where $a(t)$ and $b(t)$ obey (10-34).

If we also introduce the notation

$$\omega_1 = \frac{egB_1}{4m_e} \quad (10-36)$$

then the equations for $A(t)$ and $B(t)$ read as follows

$$\begin{aligned} i \frac{dA(t)}{dt} &= -\omega_0 A(t) + i \frac{da(t)}{dt} e^{i\omega_0 t} \\ &= -\omega_0 A(t) + (\omega_0 a(t) + \omega_1 b(t) \cos \omega t) e^{i\omega_0 t} \\ &= \omega_1 B(t) \cos \omega t e^{2i\omega_0 t} \\ &\approx \frac{1}{2} \omega_1 B(t) e^{(2i\omega_0 - \omega)t} \end{aligned} \quad (10-37)$$

In exactly the same way we obtain

$$i \frac{dB(t)}{dt} \approx \frac{1}{2} \omega_1 A(t) e^{-i(2\omega_0 - \omega)t} \quad (10-38)$$

To obtain these equations in their final form we made use of the *rotating wave* approximation

$$\cos \omega t e^{2i\omega_0 t} = \frac{1}{2} (e^{2i\omega_0 t + i\omega t} + e^{2i\omega_0 t - i\omega t}) \approx \frac{1}{2} e^{i(2\omega_0 - \omega)t} \quad (10-39)$$

The resonance condition $\omega = 2\omega_0$ implies that the first term above will oscillate very rapidly, and it contributes nothing on the average. We only used the second term in the rotating wave approximation.

The two first-order differential equations can be combined into a single second-order equation. We differentiate (10-37) and get

$$\begin{aligned} \frac{d^2 A(t)}{dt^2} &= -\frac{i\omega_1}{2} \frac{d}{dt} (e^{i(2\omega_0 - \omega)t} B(t)) \\ &= \frac{\omega_1}{2} (2\omega_0 - \omega) e^{i(2\omega_0 - \omega)t} B(t) - \frac{i\omega_1}{2} e^{i(2\omega_0 - \omega)t} \frac{dB(t)}{dt} \\ &= (2\omega_0 - \omega) i \frac{dA(t)}{dt} - \left(\frac{\omega_1}{2}\right)^2 A(t) \end{aligned} \quad (10-40)$$

To solve this differential equation, let us try the solution $A(t) = A(0)e^{i\Omega t}$. We get

$$-\Omega^2 = -(2\omega_0 - \omega)\Omega - \left(\frac{\omega_1}{2}\right)^2$$

This quadratic equation has two roots:

$$\Omega_{\pm} = \left(\omega_0 - \frac{\omega}{2}\right) \pm \sqrt{\left(\omega_0 - \frac{\omega}{2}\right)^2 + \frac{\omega_1^2}{4}} \quad (10-41)$$

We now write

$$A(t) = A_+ e^{i\Omega_+ t} + A_- e^{i\Omega_- t} \quad (10-42)$$

After a couple of lines of algebra this leads to

$$B(t) = -\frac{2}{\omega_1} (\Omega_+ A_+ e^{-i\Omega_+ t} + \Omega_- A_- e^{-i\Omega_- t}) \quad (10-43)$$

We can now finally write the solutions for $a(t)$ and $b(t)$. They are

$$\begin{aligned} a(t) &= A(t) e^{-i\omega_0 t} \\ b(t) &= B(t) e^{i\omega_0 t} \end{aligned} \quad (10-44)$$

Suppose that at time $t = 0$ the spin is in the “up” spin state χ_+ . This means that $a(0) = 1$, $b(0) = 0$, which translates into

$$\begin{aligned} A_+ + A_- &= 1 \\ \Omega_+ A_+ + \Omega_- A_- &= 0 \end{aligned} \quad (10-45)$$

The solution is

$$\begin{aligned} A_+ &= \frac{\Omega_-}{\Omega_- - \Omega_+} \\ A_- &= -\frac{\Omega_+}{\Omega_- - \Omega_+} \end{aligned} \quad (10-46)$$

We may now calculate the probability that at a time t , the system is in the “down” spin state χ_- . That probability is

$$\begin{aligned} P_-(t) &= |b(t)|^2 = \frac{4}{\omega_1^2} |\Omega_+ A_+ e^{i\Omega_+ t} + \Omega_- A_- e^{i\Omega_- t}|^2 \\ &= \frac{4}{\omega_1^2} \left(\frac{\Omega_+ \Omega_-}{\Omega_- - \Omega_+} \right)^2 |e^{i\Omega_+ t} - e^{i\Omega_- t}|^2 \\ &= \frac{8}{\omega_1^2} \left(\frac{\Omega_+ \Omega_-}{\Omega_- - \Omega_+} \right)^2 (1 - \cos(\Omega_- - \Omega_+)t) \end{aligned} \quad (10-47)$$

At resonance, when

$$\omega = 2\omega_0 \quad (10-48)$$

(10-41) yields

$$\Omega_{\pm} = \pm \frac{\omega_1}{2}$$

This means that

$$P_{\text{res}}(t) = \frac{1}{2}(1 - \cos \omega_1 t) \quad (10-49)$$

Off-resonance we have

$$P_-(t) = \frac{1}{2} \frac{\omega_1^2}{(2\omega_0 - \omega)^2 + \omega_1^2} (1 - \cos \sqrt{(2\omega_0 - \omega)^2 + \omega_1^2} t) \quad (10-50)$$

which is small, since $\omega \gg \omega_1$. At resonance the probability becomes of the order of unity. Since the energy of the “up” state is different from that of the “down” state, this energy difference, absorbed from the external field, is enhanced at resonance, when it matches $2\omega_0$. This allows us to determine g .

10-4 ADDITION OF TWO SPINS

In classical mechanics, angular momenta add vectorially. For example, if the angular momentum of the moon about the earth’s axis is $\mathbf{S}$ and the orbital angular momentum of the earth about the sun is $\mathbf{L}$, then there is meaning to the statement that the total angular momentum of the moon about the sun is

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (10-51)$$

Does this carry over to quantum mechanics? We will show that we can “add” two angular momenta, but that the addition is not vectorial in the usual sense. We begin by considering the *addition of two spins*. An example of a two-spin system might be the ground state of helium, which has two electrons, both with zero orbital angular momentum in the lowest state. Thus the only contribution to the total angular momentum comes from the spins. We ignore all other aspects of the problem and deal only with the spins.

Suppose we have two electrons, whose spins are described by the operators $\mathbf{S}_1$ and $\mathbf{S}_2$. Each of these sets of operators satisfies the standard angular momentum commutation relations

$$[S_{1x}, S_{1y}] = i\hbar S_{1z} \quad (\text{and cycl.}) \quad (10-52)$$

and

$$[S_{2x}, S_{2y}] = i\hbar S_{2z} \quad (\text{and cycl.}) \quad (10-53)$$

but the two sets of operators commute with each other, since the degrees of freedom associated with different particles are independent; that is,

$$[\mathbf{S}_1, \mathbf{S}_2] = 0$$

Let us now define the total spin $\mathbf{S}$ by

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \quad (10-54)$$

The commutation relations obeyed by the components of $\mathbf{S}$ are

$$\begin{aligned} [S_x, S_y] &= [S_{1x} + S_{2x}, S_{1y} + S_{2y}] \\ &= [S_{1x}, S_{1y}] + [S_{2x}, S_{2y}] \\ &= i\hbar(S_{1z} + S_{2z}) = i\hbar S_z \end{aligned} \quad (10-55)$$

and so on. We are therefore justified in calling $\mathbf{S}$ the total *spin*. We may now determine the eigenvalues and eigenfunctions of $\mathbf{S}^2$ and S_z .

The two-spin system actually has four states. If we denote the spinor of the first electron by $\chi_{\pm}^{(1)}$, so that

$$\begin{aligned} \mathbf{S}_1^2 \chi_{\pm}^{(1)} &= \frac{1}{2}(\frac{1}{2} + 1) \hbar^2 \chi_{\pm}^{(1)} \\ S_{1z} \chi_{\pm}^{(1)} &= \pm \frac{1}{2} \hbar \chi_{\pm}^{(1)} \end{aligned} \quad (10-56)$$

and similarly for the spinor $\chi_{\pm}^{(2)}$ of the second electron, then the four states are

$$\chi_+^{(1)} \chi_+^{(2)}, \quad \chi_+^{(1)} \chi_-^{(2)}, \quad \chi_-^{(1)} \chi_+^{(2)}, \quad \chi_-^{(1)} \chi_-^{(2)} \quad (10-57)$$

The eigenvalues of S_z for the four states are

$$\begin{aligned} S_z \chi_{\pm}^{(1)} \chi_{\pm}^{(2)} &= (S_{1z} + S_{2z}) \chi_{\pm}^{(1)} \chi_{\pm}^{(2)} \\ &= (S_{1z} \chi_{\pm}^{(1)}) \chi_{\pm}^{(2)} + \chi_{\pm}^{(1)} (S_{2z} \chi_{\pm}^{(2)}) \end{aligned}$$

that is,

$$\begin{aligned} S_z \chi_+^{(1)} \chi_+^{(2)} &= \hbar \chi_+^{(1)} \chi_+^{(2)} \\ S_z \chi_+^{(1)} \chi_-^{(2)} &= S_z \chi_-^{(1)} \chi_+^{(2)} = 0 \\ S_z \chi_-^{(1)} \chi_-^{(2)} &= -\hbar \chi_-^{(1)} \chi_-^{(2)} \end{aligned} \quad (10-58)$$

There are two states with m -value 0. One might expect that one linear combination of them will form an $S = 1$ state, to form a triplet with the $m = 1$ and $m = -1$ states, and the orthogonal combination will form a singlet $S = 0$ state. To check this expectation, let us construct the lowering operator

$$S_- = S_{1-} + S_{2-} \quad (10-59)$$

and apply this to the $m = 1$ state. This should give us the $m = 0$ state that belongs to the $S = 1$ triplet, aside from a coefficient in front. Indeed, using the fact that

$$S_-^{(i)} \chi_{\pm}^{(i)} = \hbar \chi_{\mp}^{(i)} \quad (10-60)$$

which can be established by noting that

$$\frac{1}{2}\hbar \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - i \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \right] \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \hbar \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (10-61)$$

we get

$$\begin{aligned} S_- \chi_+^{(1)} \chi_+^{(2)} &= (S_{1-} \chi_+^{(1)}) \chi_+^{(2)} + \chi_+^{(1)} S_{2-} \chi_+^{(2)} \\ &= \hbar \chi_-^{(1)} \chi_+^{(2)} + \hbar \chi_+^{(1)} \chi_-^{(2)} \\ &= \sqrt{2}\hbar \frac{\chi_+^{(1)} \chi_-^{(2)} + \chi_-^{(1)} \chi_+^{(2)}}{\sqrt{2}} \end{aligned} \quad (10-62)$$

The linear combination has been normalized, and the compensating factor in front, $\sqrt{2}\hbar$, agrees with what one would expect from (7-20) and (7-24) with $l = m = 1$. If we now apply S_- to this linear combination, and note that

$$S_-^{(0)}\chi_-^{(0)} = 0 \quad (10-63)$$

we get

$$\begin{aligned} S_- \frac{\chi_+^{(1)}\chi_-^{(2)} + \chi_-^{(1)}\chi_+^{(2)}}{\sqrt{2}} &= \frac{\hbar}{\sqrt{2}} (\chi_-^{(1)}\chi_-^{(2)} + \chi_-^{(1)}\chi_-^{(2)}) \\ &= \sqrt{2}\hbar\chi_-^{(1)}\chi_-^{(2)} \end{aligned} \quad (10-64)$$

as we should, for an angular momentum state $S = 1$. The remaining state, constructed to be orthogonal to (10-62) and properly normalized, is

$$\frac{1}{\sqrt{2}} (\chi_+^{(1)}\chi_-^{(2)} - \chi_-^{(1)}\chi_+^{(2)}) \quad (10-65)$$

and because it has no partners, we conjecture that it is an $S = 0$ state. In order to check this, we compute S^2 for the two states

$$X_{\pm} = \frac{1}{\sqrt{2}} (\chi_+^{(1)}\chi_-^{(2)} \pm \chi_-^{(1)}\chi_+^{(2)}) \quad (10-66)$$

We have

$$\begin{aligned} S^2 &= (S_1 + S_2)^2 = S_1^2 + S_2^2 + 2S_1 \cdot S_2 \\ &= S_1^2 + S_2^2 + 2S_{1z}S_{2z} + S_{1+}S_{2-} + S_{1-}S_{2+} \end{aligned} \quad (10-67)$$

First,

$$\begin{aligned} S_1^2 X_{\pm} &= \frac{1}{\sqrt{2}} (\chi_-^{(2)} S_1^2 \chi_+^{(1)} \pm \chi_+^{(2)} S_1^2 \chi_-^{(1)}) \\ &= \frac{3}{4}\hbar^2 X_{\pm} \end{aligned} \quad (10-68)$$

and similarly,

$$S_2^2 X_{\pm} = \frac{3}{4}\hbar^2 X_{\pm} \quad (10-69)$$

Next, we calculate

$$2S_{1z}S_{2z}X_{\pm} = 2(\frac{1}{2}\hbar)(-\frac{1}{2}\hbar)X_{\pm} = -\frac{1}{2}\hbar^2 X_{\pm} \quad (10-70)$$

Finally,

$$\begin{aligned} (S_{1+}S_{2-} + S_{1-}S_{2+})X_{\pm} &= \frac{1}{\sqrt{2}} (S_{1+}\chi_+^{(1)}S_{2-}\chi_-^{(2)} + S_{1-}\chi_+^{(1)}S_{2+}\chi_-^{(2)} \\ &\quad \pm S_{1+}\chi_-^{(1)}S_{2-}\chi_+^{(2)} \pm S_{1-}\chi_-^{(1)}S_{2+}\chi_+^{(2)}) \end{aligned}$$

which, with the help of (10-60) and (10-63), yields

$$(S_{1+}S_{2-} + S_{1-}S_{2+})X_{\pm} = \pm\hbar X_{\pm} \quad (10-71)$$

Thus

$$\begin{aligned} S^2 X_{\pm} &= \hbar^2(\frac{3}{4} + \frac{3}{4} - \frac{1}{2} \pm 1) X_{\pm} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \hbar^2 X_{\pm} \\ &= \hbar^2 S(S+1) X_{\pm} \end{aligned} \quad (10-72)$$

with $S = 1$ and 0 corresponding to the $\pm$ states.

What we have shown is that the totality of the four states of two spin 1/2 particles may be recombined into a triplet and into a singlet total spin state. It is important to note that the two descriptions are entirely equivalent. In one case we have as our complete set of commuting observables S_1^2 , S_2^2 , S_{1z} , and S_{2z} . In the other case we have as our complete set of commuting observables S^2 , S_z , S_1^2 , S_2^2 . By the expansion theorem, any function can be expanded in terms of a complete set of eigenstates. *What we have demonstrated here is the expansion of the eigenstates of the second set of observables in terms of the complete set of states of the first set of observables.* This is quite analogous to the expression of the eigenstates of the hydrogen atom in terms of the eigenstates of the momentum operator, in which the coefficients (the analogs of the $1/\sqrt{2}$'s here) are the momentum-space wave functions. It is a simple exercise to invert the process and to find the products of the $\chi^{(1)}\chi^{(2)}$ in terms of triplet and singlet combinations.

In physical problems it frequently happens that to first approximation two sets of completely commuting observables are equally useful in the construction of eigenstates. In the next approximation, when additional terms in the Hamiltonian are taken into account, only one of these sets remains useful. A simple example occurs in low-energy nuclear physics.

In early studies of the potential $V(r)$ that describes the interaction between neutrons and protons at low energies, it became clear that the strength of the interaction depended on whether the two interacting particles were in a total spin $S = 1$ state or a spin $S = 0$ state. For example, a deuteron had $S = 1$, while the corresponding $S = 0$ state of a neutron and a proton did not bind. This can be described in terms of a spin-dependent potential. Suppose we have

$$V(r) = V_1(r) + \frac{1}{\hbar^2} \mathbf{S}_1 \cdot \mathbf{S}_2 V_2(r) \quad (10-73)$$

we can easily see that S_{1z} and S_{2z} do not commute with the second term, so that the eigenstates of H containing this potential cannot just be simple products of eigenstates of S_{1z} and S_{2z} . If we observe, however, that

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = \frac{1}{2}(\mathbf{S}^2 - \mathbf{S}_1^2 - \mathbf{S}_2^2) \quad (10-74)$$

so that this term can be replaced by the eigenvalue, when acting on an eigenfunction of $\mathbf{S}^2$, $\mathbf{S}_1^2$, and $\mathbf{S}_2^2$, then

$$\begin{aligned} V(r) &= V_1(r) + \frac{1}{2} V_2(r) \left[S(S+1) - \frac{3}{2} \right] \\ &= V_1(r) + \frac{1}{4} \begin{pmatrix} 1 \\ -3 \end{pmatrix} V_2(r) \begin{cases} S=1 \\ S=0 \end{cases} \end{aligned} \quad (10-75)$$

Such a spin-dependent potential is actually observed in the neutron-proton system. The bound state is an $S = 1$ state—this is the deuteron—but there is also an unbound $S = 0$ state. This implies that $V_1 - (3/4)V_2$ is a less attractive potential than $V_1 + (1/4)V_2$. This is only possible if $V_2(r) \neq 0$.

The spin singlet wave function (10-65) implies that if in a measurement electron (1) is found in a spin “up” state, then electron (2) *must* be in a spin “down” state. Consider a singlet state in which the electrons are moving to the right and left, respectively, with equal and opposite momenta, so that the two-electron system is still at rest in the center of mass. We may *define* the right-moving electron as labeled with (1) and the left-moving electron as labeled with (2). Suppose we now make a measurement of the S_x component of the right-moving electron. From what we have learned so far, the eigenvalues $\pm\hbar/2$

will turn up with equal probabilities. Suppose that a particular measurement yields the (+) eigenvalue. This projects the eigenstate of electron (1) into the state $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. One might expect that a simultaneous measurement of S_x carried out on electron (2) is unpredictable, leading to the eigenvalues $\pm\hbar/2$ with equal probability, on the ground that the information about the state of electron (1) cannot propagate to the apparatus measuring electron (2) with infinite speed. *This is false!* In quantum mechanics, the two-electron system is described by a single state vector in which the spins are *entangled*. The system is still a singlet system, and the measurement of S_x must yield the (-) eigenvalue. We can see this directly. The singlet state has the wave function

$$\psi(1, 2) = \frac{1}{\sqrt{2}} \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 - \begin{pmatrix} 0 \\ 1 \end{pmatrix}_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right) \quad (10-76)$$

We now project the state of electron (1) onto the S_x eigenstate with eigenvalue $+\hbar/2$. The projection operator for the positive eigenvalue of S_x for electron (1) is given by

$$P_1(+) = \xi_+^{(1)} \xi_+^{(1)*} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{1}{\sqrt{2}} (1 \quad 1) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (10-77)$$

When this acts on the (1) component of the state (10-76) we get

$$\begin{aligned} P_1(+)\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 \right\} - P_1(+)\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}_1 \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right\}\right) \\ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 - \frac{1}{\sqrt{2}} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 = -\frac{1}{\sqrt{2}} \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_1 \times \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}_2 \right\} \end{aligned} \quad (10-78)$$

This, aside from the numerical factor $-1/\sqrt{2}$ in front, is just the product that describes electron (1) in an eigenstate of S_x with eigenvalue $\hbar/2$ and electron (2) in an eigenstate of S_x with eigenvalue $-\hbar/2$, as expected as a result of the *entanglement* of the two spins.

Another way of seeing this is to expand the spinor eigenstates $\chi_{\pm}$ of S_z in terms of the spinor eigenstates $\xi_{\pm}$ of S_x . We have [cf. (10-17) and (10-18) with $\phi = 0$]

$$\xi_{\pm} = \frac{1}{\sqrt{2}} (\chi_{+} \pm \chi_{-})$$

or, equivalently

$$\chi_{\pm} = \frac{1}{\sqrt{2}} (\xi_{+} \pm \xi_{-})$$

If we substitute these into

$$\psi = \frac{1}{\sqrt{2}} (\chi_{+}^{(1)} \chi_{-}^{(2)} - \chi_{-}^{(1)} \chi_{+}^{(2)})$$

we get, after a little algebra,

$$\psi = \frac{1}{\sqrt{2}} (\xi_{+}^{(1)} \xi_{-}^{(2)} - \xi_{-}^{(1)} \xi_{+}^{(2)})$$

which clearly shows the correlation of spins in the x -direction.

10-5 THE ADDITION OF SPIN 1/2 AND ORBITAL ANGULAR MOMENTUM

The addition of a spin angular momentum with an orbital angular momentum is of great importance in atomic and nuclear physics. The procedure for doing this is exactly the same as the one we used to add two spins. The algebraic details are worked out in Supplement 10-A [www.wiley.com/college/gasiorowicz], where we also give a superficial discussion of the addition of two arbitrary angular momenta. Here, we will just write down the results.

For

$$\mathbf{J} = \mathbf{S} + \mathbf{L} \quad (10-79)$$

we can find linear combinations of the spin eigenstates $\chi_{\pm}$ and the orbital angular momentum eigenstates $Y_{lm}(\theta, \varphi)$ that are eigenstates of the total angular momentum. We find

$$\begin{aligned} \mathbf{J}^2 \psi_{j,m_j} &= \hbar^2 j(j+1) \psi_{j,m_j} \\ J_z \psi_{j,m_j} &= \hbar m_j \psi_{j,m_j} \end{aligned} \quad (10-80)$$

where j can take on the values $j = l \pm \frac{1}{2}$. In both cases,

$$m_j = j, j-1, j-2, \dots, -j \quad (10-81)$$

The suitably normalized linear combinations are

$$\begin{aligned} \psi_{l+1/2, m+1/2} &= \sqrt{\frac{l+m+1}{2l+1}} Y_{l,m} \chi_+ + \sqrt{\frac{l-m}{2l+1}} Y_{l,m+1} \chi_- \\ \psi_{l-1/2, m+1/2} &= \sqrt{\frac{l-m}{2l+1}} Y_{l,m} \chi_+ - \sqrt{\frac{l+m+1}{2l+1}} Y_{l,m+1} \chi_- \end{aligned} \quad (10-82)$$

These eigenfunctions are orthogonal to each other.

General Rules for Addition of Angular Momenta

For $\mathbf{J} = \mathbf{L}_1 + \mathbf{L}_2$ we can find linear combinations of the eigenstates $Y_{l_1 m_1}$ and the eigenstates $Y_{l_2 m_2}$ that are eigenstates of the total angular momentum. We find, as expected, that

$$\begin{aligned} \mathbf{J}^2 \psi_{j,m_j} &= \hbar^2 j(j+1) \psi_{j,m_j} \\ J_z \psi_{j,m_j} &= \hbar m_j \psi_{j,m_j} \end{aligned} \quad (10-83)$$

where j can take on the values

$$j = l_1 + l_2, l_1 + l_2 - 1, l_1 + l_2 - 2, \dots, |l_1 - l_2| \quad (10-84)$$

and

$$m_j = j, j-1, j-2, \dots, -j \quad (10-85)$$

We associated angular momentum with integer values with orbital angular momentum. This is not always so. There are particles that have an *intrinsic* spin 1, just as electrons have an intrinsic spin 1/2. The states of these particles in isolation will be described by matrices, except that these are now 3×3 matrices. We saw these in equations (9-34), (9-37), and (9-38). For an interesting digression on this subject, see Supplement 10-B [www.wiley.com/college/gasiorowicz].

PROBLEMS

1. Find the normalized eigenspinors for S_y .
2. Find the normalized eigenspinors of the 2×2 matrix

$$\begin{pmatrix} \cos \alpha & \sin \alpha e^{-i\beta} \\ \sin \alpha e^{i\beta} & -\cos \alpha \end{pmatrix}$$

3. Use the method outlined in Chapter 9 to find the unitary matrix that diagonalizes the above matrix.
4. Use material from Chapter 9 to construct the matrix representations of spin $3/2$. Find the eigenstates in the representation in which S_z is diagonal.
5. Consider the spinor $\frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. What is the probability that a measurement of $(3S_x + 4S_y)/5$ yields the value $-\hbar/2$?
6. Consider a spin $1/2$ system represented by the normalized spinor $\frac{1}{\sqrt{65}} \begin{pmatrix} 4 \\ 7 \end{pmatrix}$. What is the probability that a measurement of S_y yields the value $-\hbar/2$?
7. Prove that $(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i \boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B})$
8. A spin $1/2$ particle is in an eigenstate of S_x with eigenvalue $\hbar/2$ at time $t = 0$. At that time it is placed in a magnetic field of magnitude B pointing in the z -direction, in which it is allowed to precess for a time T . At that instant, the magnetic field is rotated *very rapidly*, so that it now points in the y -direction. After another time interval T , a measurement of S_x is carried. What is the probability that the value $\hbar/2$ will be found?
9. (a) Show that any 2×2 matrix may be written in the form

$$A + \mathbf{B} \cdot \boldsymbol{\sigma}$$

(b) What are the conditions that A and $\mathbf{B}$ must satisfy if the matrix is to be unitary? If it is to be hermitian?

10. Show that for any vector $\mathbf{a}$ with a magnitude a ,

$$e^{i\boldsymbol{\sigma} \cdot \mathbf{a}} = \cos a + i \boldsymbol{\sigma} \cdot \mathbf{a} \frac{\sin a}{a}$$

11. Consider two spin $1/2$ particles, whose spins are described by the Pauli operators $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$. Let $\hat{\mathbf{e}}$ be the unit vector connecting the two particles and define the operator

$$S_{12} = 3(\boldsymbol{\sigma}_1 \cdot \hat{\mathbf{e}})(\boldsymbol{\sigma}_2 \cdot \hat{\mathbf{e}}) - \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2$$

Show that if the two particles are in an $S = 0$ state (singlet) then

$$S_{12} X_{\text{singlet}} = 0$$

Show that for a triplet state

$$(S_{12} - 2)(S_{12} + 4) X_{\text{triplet}} = 0$$

(Hint: Choose $\hat{\mathbf{e}}$ along the z -axis.)

12. In a low energy neutron–proton system (which has zero orbital angular momentum) the potential energy is given by

$$V(r) = V_1(r) + V_2(r) \left(3 \frac{(\boldsymbol{\sigma}_1 \cdot \mathbf{r})(\boldsymbol{\sigma}_2 \cdot \mathbf{r})}{r^2} - \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 \right) + V_3(r) \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2$$

where $\mathbf{r}$ is the vector connecting the two particles. Calculate the potential energy for the neutron–proton system (a) In the spin singlet state, and (b) In the triplet state.

13. Consider two electrons in a spin singlet state.

(a) If a measurement of the spin of one of the electrons shows that it is in a state with $s_z = 1/2$, what is the probability that a measurement of the z -component of the spin of the other electron yields $s_z = 1/2$?

(b) If a measurement of the spin of one of the electrons shows that it is in a state with $s_y = 1/2$, what is the probability that a measurement of the x -component of the spin yields $s_x = 1/2$ for the second electron?

(c) If electron (1) is in a state described by $\cos \alpha_1 \chi_+ + \sin \alpha_1 e^{i\beta_1} \chi_-$ and electron (2) is in a state described by $\cos \alpha_2 \chi_+ + \sin \alpha_2 e^{i\beta_2} \chi_-$, what is the probability that the two-electron state is in a triplet state?

14. A particle of spin 1 moves in a central potential of the form

$$V(r) = V_1(r) + \frac{\mathbf{S} \cdot \mathbf{L}}{\hbar^2} V_2(r) + \frac{(\mathbf{S} \cdot \mathbf{L})^2}{\hbar^4} V_3(r)$$

What are the values of $V(r)$ in the states $J = L + 1$, L , and $L - 1$?

Supplement 10-A

The Addition of Spin 1/2 and Orbital Angular Momentum (Details)

Of great importance for future applications is the combination of a spin with an orbital angular momentum. Since $\mathbf{L}$ depends on spatial coordinates and $\mathbf{S}$ does not, they commute

$$[\mathbf{L}, \mathbf{S}] = 0 \quad (10A-1)$$

It is therefore evident that the components of the total angular momentum $\mathbf{J}$, defined by

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (10A-2)$$

will satisfy the angular momentum commutation relations.

In asking for linear combinations of the Y_{lm} and the $\chi_{\pm}$ that are eigenstates of

$$J_z = L_z + S_z \quad (10A-3)$$

and

$$\begin{aligned} \mathbf{J}^2 &= \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L} \cdot \mathbf{S} \\ &= \mathbf{L}^2 + \mathbf{S}^2 + 2L_z S_z + L_+ S_- + L_- S_+ \end{aligned} \quad (10A-4)$$

we are again looking for the expansion coefficients of one complete set of eigenfunctions in terms of another set of eigenfunctions.

Let us consider the linear combination

$$\psi_{j,m+1/2} = \alpha Y_{lm} \chi_+ + \beta Y_{l,m+1} \chi_- \quad (10A-5)$$

It is, by construction, an eigenfunction of J_z with eigenvalue $(m + \frac{1}{2})\hbar$. We now determine α and β such that it is also an eigenfunction of J^2 . We shall make use of the fact that

$$\begin{aligned} L_+ Y_{lm} &= [l(l+1) - m(m+1)]^{1/2} \hbar Y_{l,m+1} \\ &= [(l+m+1)(l-m)]^{1/2} \hbar Y_{l,m+1} \\ L_- Y_{lm} &= [(l-m+1)(l+m)]^{1/2} \hbar Y_{l,m-1} \\ S_+ \chi_+ &= S_- \chi_- = 0 \quad S_{\pm} \chi_{\mp} = \hbar \chi_{\pm} \end{aligned} \quad (10A-6)$$

Then

$$\begin{aligned} J^2 \psi_{j,m+1/2} &= \alpha \hbar^2 \left\{ l(l+1) Y_{lm} \chi_+ + \frac{3}{4} Y_{lm} \chi_+ + 2m(\frac{1}{2}) Y_{lm} \chi_+ \right. \\ &\quad \left. + [(l-m)(l+m+1)]^{1/2} Y_{l,m+1} \chi_- \right\} + \beta \hbar^2 \left\{ l(l+1) Y_{l,m+1} \chi_- \right. \\ &\quad \left. + \frac{3}{4} Y_{l,m+1} \chi_- + 2(m+1)(-\frac{1}{2}) Y_{l,m+1} \chi_- \right. \\ &\quad \left. + [(l-m)(l+m+1)]^{1/2} Y_{lm} \chi_+ \right\} \end{aligned} \quad (10A-7)$$

This will be of the form

$$\hbar^2 j(j+1) \psi_{j,m+1/2} = \hbar^2 j(j+1) (\alpha Y_{lm} \chi_+ + \beta Y_{l,m+1} \chi_-) \quad (10A-8)$$

provided that

$$\begin{aligned} \alpha[l(l+1) + \frac{3}{4} + m] + \beta[(l-m)(l+m+1)]^{1/2} &= j(j+1) \alpha \\ \beta[l(l+1) + \frac{3}{4} - m - 1] + \alpha[(l-m)(l+m+1)]^{1/2} &= j(j+1) \beta \end{aligned} \quad (10A-9)$$

This requires that

$$\begin{aligned} (l-m)(l+m+1) &= [j(j+1) - l(l+1) - \frac{3}{4} - m] \\ &\times [j(j+1) - l(l+1) - \frac{3}{4} + m + 1] \end{aligned}$$

which evidently has two solutions,

$$j(j+1) - l(l+1) - \frac{3}{4} = \begin{cases} -l-1 \\ l \end{cases} \quad (10A-10)$$

that is,

$$j = \begin{cases} l - \frac{1}{2} \\ l + \frac{1}{2} \end{cases} \quad (10A-11)$$

For $j = l + 1/2$, we get, after a little algebra

$$\alpha = \sqrt{\frac{l+m+1}{2l+1}} \quad \beta = \sqrt{\frac{l-m}{2l+1}} \quad (10A-12)$$

(Actually we just get the ratio; these are already normalized forms.) Thus

$$\psi_{l+1/2,m+1/2} = \sqrt{\frac{l+m+1}{2l+1}} Y_{lm} \chi_+ + \sqrt{\frac{l-m}{2l+1}} Y_{l,m+1} \chi_- \quad (10A-13)$$

We can guess that the $j = l - 1/2$ solution must have the form

$$\psi_{l-1/2,m+1/2} = \sqrt{\frac{l-m}{2l+1}} Y_{lm} \chi_+ - \sqrt{\frac{l+m+1}{2l+1}} Y_{l,m+1} \chi_- \quad (10A-14)$$

in order to be orthogonal to the $j = l + 1/2$ solution.

General Rules for Addition of Angular Momenta, and Implications for Identical Particles

These two examples illustrate the general features that are involved in the addition of angular momenta: If we have the eigenstates $Y_{l_1 m_1}^{(1)}$ of $\mathbf{L}_1^2$ and L_{1z} , and the eigenstates $Y_{l_2 m_2}^{(2)}$ of $\mathbf{L}_2^2$ and L_{2z} , then we can form $(2l_1 + 1)(2l_2 + 1)$ product wave functions

$$Y_{l_1 m_1}^{(1)} Y_{l_2 m_2}^{(2)} \left\{ \begin{array}{l} -l_1 \leq m_1 \leq l_1 \\ -l_2 \leq m_2 \leq l_2 \end{array} \right\} \quad (10A-15)$$

These can be classified by the eigenvalue of

$$J_z = L_{1z} + L_{2z} \quad (10A-16)$$

which is $m_1 + m_2$, and which ranges from a maximum value of $l_1 + l_2$ down to $-l_1 - l_2$. As in the simple cases discussed earlier, different linear combinations of functions with

the same m value will belong to different values of j . In the following table we list the possible combinations for the special example of $l_1 = 4, l_2 = 2$. We shall use the simple abbreviation (m_1, m_2) for $Y_{l_1 m_1}^{(1)} Y_{l_2 m_2}^{(2)}$.

m -value	m_1, m_2 combinations	numbers
6	(4, 2)	1
5	(4, 1) (3, 2)	2
4	(4, 0) (3, 1) (2, 2)	3
3	(4, -1) (3, 0) (2, 1) (1, 2)	4
2	(4, -2) (3, -1) (2, 0) (1, 1) (0, 2)	5
1	(3, -2) (2, -1) (1, 0) (0, 1) (-1, 2)	5
0	(2, -2) (1, -1) (0, 0) (-1, 1) (-2, 2)	5
-1	(1, -2) (0, -1) (-1, 0) (-2, 1) (-3, 2)	5
-2	(0, -2) (-1, -1) (-2, 0) (-3, 1) (-4, 2)	5
-3	(-1, -2) (-2, -1) (-3, 0) (-4, 1)	4
-4	(-2, -2) (-3, -1) (-4, 0)	3
-5	(-3, -2) (-4, -1)	2
-6	(-4, -2)	1

There are a total of 45 combinations, consistent with $(2l_1 + 1)(2l_2 + 1)$.

The highest state has total angular momentum $l_1 + l_2$ as can easily be checked by applying J^2 to $Y_{l_1 l_1}^{(1)} Y_{l_2 l_2}^{(2)}$:

$$\begin{aligned}
 \mathbf{J}^2 Y_{l_1 l_1}^{(1)} Y_{l_2 l_2}^{(2)} &= (\mathbf{L}_1^2 + \mathbf{L}_2^2 + 2L_{1z}L_{2z} + L_{1+}L_{2-} + L_{1-}L_{2+}) Y_{l_1 l_1}^{(1)} Y_{l_2 l_2}^{(2)} \\
 &= \hbar^2 [l_1(l_1 + 1) + l_2(l_2 + 1) + 2l_1 l_2] Y_{l_1 l_1}^{(1)} Y_{l_2 l_2}^{(2)} \\
 &= \hbar^2 (l_1 + l_2)(l_1 + l_2 + 1) Y_{l_1 l_1}^{(1)} Y_{l_2 l_2}^{(2)}
 \end{aligned} \tag{10A-17}$$

This is $j = 6$ in the example discussed in the table. Successive applications of

$$J_- = L_{1-} + L_{2-} \tag{10A-18}$$

will pick out one linear combination from each row in the table. These will form the 13 states that belong to $j = 6$. When this is done, there remains a single state with $m = 5$, two with $m = 4, \dots$, one with $m = -5$. It is extremely plausible, and can, in fact, be checked, that the $m = 5$ state belongs to $j = 5$. Again, successive applications of J_- pick out another linear combination from each row in the table, forming 11 states that belong to $j = 5$. Repetition of this procedure shows that we get, after this, sets that belong to $j = 4, j = 3$, and finally $j = 2$. The multiplicities add up to 45:

$$13 + 11 + 9 + 7 + 5 = 45$$

We shall not work out the details of this decomposition, as it is beyond the scope of this book. We merely state the results.

- (a) The product $Y_{l_1 m_1}^{(1)} Y_{l_2 m_2}^{(2)}$ can be decomposed into eigenstates of $\mathbf{J}^2$, with eigenvalues $j(j + 1)\hbar^2$, where j can take on the values

$$j = l_1 + l_2, l_1 + l_2 - 1, \dots, |l_1 - l_2| \tag{10A-19}$$

We can verify that the multiplicities check in (10A-19): If we sum the number of states, we get ($l_1 \geq l_2$)

$$\begin{aligned} & [2(l_1 + l_2) + 1] + [2(l_1 + l_2 - 1) + 1] + \cdots + [2(l_1 - l_2) + 1] \\ &= \sum_{n=0}^{2l_2} [2(l_1 - l_2 + n) + 1] \\ &= (2l_2 + 1)(2l_1 + 1) \end{aligned} \quad (10A-20)$$

- (b) It is possible to generatize (10A-13) and (10A-14) to give the Clebsch-Gordan series

$$\psi_{jm} = \sum_{m_1} C(jm; l_1 m_1 l_2 m_2) Y_{l_1 m_1}^{(1)} Y_{l_2 m_2}^{(2)} \quad (10A-21)$$

The coefficients $C(jm; l_1 m_1 l_2 m_2)$ are called Clebsch-Gordan coefficients, and they have been tabulated for many values of the arguments. We calculated the coefficients for $l_2 = 1/2$, and summarize (10A-12) and (10A-13) in the table that follows. Note that $m = m_1 + m_2$, so that the m in (10A-13) and (10A-14) is really m_1 below.

$$C(jm; l_1 m_1, 1/2, m_2)$$

	$m_2 = 1/2$	$m_2 = -1/2$
$j = l_1 + 1/2$	$\sqrt{\frac{l_1 + m + 1/2}{2l_1 + 1}}$	$\sqrt{\frac{l_1 - m + 1/2}{2l_1 + 1}}$
$j = l_1 - 1/2$	$-\sqrt{\frac{l_1 - m + 1/2}{2l_1 + 1}}$	$\sqrt{\frac{l_1 + m + 1/2}{2l_1 + 1}}$

Another useful table is

$$C(jm; l_1 m_1, 1, m_2)$$

	$m_2 = 1$	$m_2 = 0$	$m_2 = -1$
$j = l_1 + 1$	$\sqrt{\frac{(l_1 + m)(l_1 + m + 1)}{(2l_1 + 1)(2l_1 + 2)}}$	$\sqrt{\frac{(l_1 - m + 1)(l_1 + m + 1)}{(2l_1 + 1)(l_1 + 1)}}$	$\sqrt{\frac{(l_1 - m)(l_1 - m + 1)}{(2l_1 + 1)(2l_1 + 2)}}$
$j = l_1$	$-\sqrt{\frac{(l_1 + m)(l_1 - m + 1)}{2l_1(l_1 + 1)}}$	$\frac{m}{\sqrt{l_1(l_1 + 1)}}$	$\sqrt{\frac{(l_1 - m)(l_1 + m + 1)}{2l_1(2l_1 + 1)}}$
$j = l_1 - 1$	$\sqrt{\frac{(l_1 - m)(l_1 - m + 1)}{2l_1(2l_1 + 1)}}$	$-\sqrt{\frac{(l_1 - m)(l_1 + m)}{l_1(2l_1 + 1)}}$	$\sqrt{\frac{(l_1 + m)(l_1 + m + 1)}{2l_1(2l_1 + 1)}}$

Supplement 10-B

The Levi-Civita Symbol and Maxwell's Equations

A very useful mathematical device is the use of the *Levi-Civita* symbol. The symbol e_{ijk} is defined by the following properties:

- (a) It is antisymmetric under the interchange of any two of its indices. For example,

$$e_{123} = -e_{213} = -(-e_{231}) \quad (10B-1)$$

and so on. Two consequences of this rule are

- (i) When any two indices are equal, the value of e_{ijk} is zero.

- (ii) $e_{123} = e_{231} = e_{312}$

(b) $e_{123} = 1$ (10B-2)

Some consequences of this definition are

$$\begin{aligned} e_{ijk}e_{ijm} &= 2\delta_{km} \\ e_{ijk}e_{imn} &= \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km} \\ e_{ijk}A_jB_k &= (A \times B)_i \\ [L_i, L_j] &= ie_{ijk}L_k \end{aligned} \quad (10B-3)$$

We may use this to write out Maxwell's equations in a particularly interesting way.

Maxwell's equations in empty space have the form

$$\begin{aligned} \nabla \cdot \mathbf{B} &= 0 \\ \nabla \cdot \mathbf{E} &= 0 \\ \nabla \times \mathbf{B} &= \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \end{aligned} \quad (10B-4)$$

They may be rewritten in the form

$$\frac{\partial}{\partial t}(\mathbf{E} + ic\mathbf{B}) = -ic\nabla \times (\mathbf{E} + ic\mathbf{B}) \quad (10B-5)$$

Which bears some resemblance to the Schrödinger equation in that the “wave function” is complex, and that the first-order time derivative enters into the equation.

We may write the equation in a very suggestive way by using the Levi-Civita symbol in two contexts. First, the symbol may be used to give a matrix representation of the spin 1 angular momentum $\mathbf{S}$. (We are actually working with the angular momentum matrix divided by $\hbar$ —that is, with the analog of $\sigma/2$.)

To see this, we postulate

$$(S_i)_{jk} = -ie_{ijk} \quad (10B-6)$$

The square of the matrix is easily calculated. We have

$$(S^2)_{jl} = (S_i)_{jk}(S_i)_{kl} = -e_{ijk}e_{ikl} = e_{ijk}e_{ikl} = 2\delta_{jl} \quad (10B-7)$$

We next need to check the commutation relations

$$\begin{aligned} (S_a)_{jk}(S_b)_{kl} - (S_b)_{jk}(S_a)_{kl} &= -e_{ajk}e_{bkl} + e_{bjk}e_{akl} \\ &= e_{ajk}e_{bkl} - e_{bjk}e_{alk} = \delta_{ab}\delta_{jl} - \delta_{al}\delta_{jb} - \delta_{ab}\delta_{jl} + \delta_{bl}\delta_{aj} \\ &= \delta_{bl}\delta_{aj} - \delta_{al}\delta_{jb} = e_{bam}e_{ljm} = ie_{abm}(-ie_{mjl}) = ie_{abm}(S_m)_{jl} \end{aligned}$$

that is,

$$[S_a, S_b] = ie_{abm}S_m \quad (10B-8)$$

Let us now rewrite our version of Maxwell's equations. It reads

$$\begin{aligned} \frac{\partial}{\partial t}(E_i + icB_i) &= -ice_{imn}\frac{\partial}{\partial x_m}(E_n + iB_n) \\ &= -c(S_m)_{in}\frac{\partial}{\partial x_m}(E_n + iB_n) \end{aligned}$$

or equivalently,

$$i\hbar\frac{\partial}{\partial t}(E_i + icB_i) = c(S_m)_{in}\frac{\hbar}{i}\frac{\partial}{\partial x_m}(E_n + iB_n) \quad (10B-9)$$

With the notation $\psi_i = (E_i + icB_i)$, we get

$$i\hbar\frac{\partial\psi_i}{\partial t} = c(\mathbf{S} \cdot \mathbf{p}_{\text{op}})_{im}\psi_m \quad (10B-10)$$

The operator on the right side is the projection of the photon spin along the direction of motion. The complex conjugate wave function is easily seen to satisfy

$$i\hbar\frac{\partial\psi_i^*}{\partial t} = -c(\mathbf{S} \cdot \mathbf{p}_{\text{op}})_{im}\psi_m^* \quad (10B-11)$$

where the right side represents the opposite projection (helicity). We need both equations to obtain separate equations for $\mathbf{E}$ and $\mathbf{B}$.

Chapter 11

Time-Independent Perturbation Theory

11-1 ENERGY SHIFTS AND PERTURBED EIGENSTATES

There are few potentials $V(r)$ for which the Schrödinger equation is exactly solvable, and we have already discussed most of them. We must therefore develop techniques for obtaining eigenvalues and eigenfunctions for all the other potentials. The general availability of computers would seem to render this effort as unnecessary, and indeed, if only numerical values of energy eigenvalues, for example, are needed, the techniques that we develop in this chapter cannot compete in accuracy with the simplest programs. The reason for discussing perturbation theory is that even at the lowest level of approximation we get insight into how energies shift and how eigenfunctions are changed by a change in the potential.

Let us assume that we have found the eigenvalues and the complete set of normalized eigenfunctions for a Hamiltonian H_0 ,

$$H_0|\phi_n\rangle = E_n^{(0)}|\phi_n\rangle \quad (11-1)$$

The wave functions in x -space will be described by $\phi_n(\mathbf{r}) = \langle \mathbf{r} | \phi_n \rangle$, but at this stage we will stay with the general formalism. We now ask to solve

$$(H_0 + \lambda H_1)|\psi_n\rangle = E_n|\psi_n\rangle \quad (11-2)$$

In general, we may view the Hamiltonian H as a matrix. We can write

$$H_{mn} \equiv \langle \phi_m | H | \phi_n \rangle = E_n^{(0)} \delta_{mn} + \langle \phi_m | \lambda H_1 | \phi_n \rangle$$

For a finite-dimensional space; that is, if the matrix is finite dimensional, diagonalization of the matrix is equivalent to solving a set of N algebraic equations. Cases of 2×2 and 3×3 matrices will be discussed in Chapter 18. In general, however, the matrices are infinite dimensional, and there we turn to perturbation theory.

We will express the desired quantities as power series in λ . The question of convergence of the series will not be discussed. Frequently one can show that the series cannot be convergent, and yet the first few terms, when λ is small, do properly describe the physical system. We will assume that as $\lambda \rightarrow 0$, $E_n \rightarrow E_n^{(0)}$.

Since the $|\phi_n\rangle$ form a complete set, we may expand the $|\psi_n\rangle$ in a series involving the $|\phi_n\rangle$. We write

$$|\psi_n\rangle = N(\lambda) \left\{ |\phi_n\rangle + \sum_{k \neq n} C_{nk}(\lambda) |\phi_k\rangle \right\} \quad (11-3)$$

The factor $N(\lambda)$ will be determined by the condition $\langle \psi_n | \psi_n \rangle = 1$. We have the freedom to choose the phase of $|\psi_n\rangle$, and we choose it such that the coefficient of $|\phi_n\rangle$ in the expansion is (aside from the normalization factor) unity. Since we require that $|\psi_n\rangle \rightarrow |\phi_n\rangle$ as $\lambda \rightarrow 0$, we must have

$$\begin{aligned} N(0) &= 1 \\ C_{nk}(0) &= 0 \end{aligned} \quad (11-4)$$

More generally, we have

$$C_{nk}(\lambda) = \lambda C_{nk}^{(1)} + \lambda^2 C_{nk}^{(2)} + \dots \quad (11-5)$$

and

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (11-6)$$

The Schrödinger energy eigenvalue equation then reads

$$\begin{aligned} (H_0 + \lambda H_1) \left\{ |\phi_n\rangle + \lambda \sum_{k \neq n} C_{nk}^{(1)} |\phi_k\rangle + \lambda^2 \sum_{k \neq n} C_{nk}^{(2)} |\phi_k\rangle + \dots \right\} \\ = (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots) \left\{ |\phi_n\rangle + \lambda \sum_{k \neq n} C_{nk}^{(1)} |\phi_k\rangle + \lambda^2 \sum_{k \neq n} C_{nk}^{(2)} |\phi_k\rangle + \dots \right\} \end{aligned} \quad (11-7)$$

Note that $N(\lambda)$ does not appear in this equation. This is a consequence of the linearity of the equation. Identifying powers of λ yields a series of equations. The first one is

$$H_0 \sum_{k \neq n} C_{nk}^{(1)} |\phi_k\rangle + H_1 |\phi_n\rangle = E_n^{(0)} \sum_{k \neq n} C_{nk}^{(1)} |\phi_k\rangle + E_n^{(1)} |\phi_n\rangle \quad (11-8)$$

Using $H_0 |\phi_n\rangle = E_n^{(0)} |\phi_n\rangle$, we obtain

$$E_n^{(1)} |\phi_n\rangle = H_1 |\phi_n\rangle + \sum_{k \neq n} (E_k^{(0)} - E_n^{(0)}) C_{nk}^{(1)} |\phi_k\rangle$$

If we now take the scalar product with $\langle \phi_n |$ and use the orthonormality condition $\langle \phi_k | \phi_l \rangle = \delta_{kl}$ we obtain

$$\lambda E_n^{(1)} = \langle \phi_n | \lambda H_1 | \phi_n \rangle \quad (11-9)$$

This is a *very important formula*. It states that the first-order energy shift for a given state is just the expectation value of the perturbing potential in that state. If the change in the potential is of a definite sign, then the energy shift will have the same sign.

In x -space, we have

$$\lambda E_n^{(1)} = \int d^3 r \int d^3 r' \langle \phi_n | \mathbf{r} \rangle \langle \mathbf{r} | \lambda H_1 | \mathbf{r}' \rangle \langle \mathbf{r}' | \phi_n \rangle$$

If H_1 depends only on $\mathbf{r}$, as would be the case with a local potential, then $\langle \mathbf{r} | H_1 | \mathbf{r}' \rangle = \delta(\mathbf{r} - \mathbf{r}') H_1(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}') V(\mathbf{r})$ and we get

$$\lambda E_n^{(1)} = \int d^3 r \phi_n^*(\mathbf{r}) \lambda V(\mathbf{r}) \phi_n(\mathbf{r}) \quad (11-10)$$

The equation shows that for the shift to be significant, the region in which the perturbing potential $V(\mathbf{r})$ is important and the region in which the probability density $|\phi_n(\mathbf{r})|^2$ is large must overlap. Note that the relation (11-9) is not tied to x -space. It applies equally well to matrices.

If we take the scalar product of (11-8) with $\langle \phi_m |$ for $m \neq n$, then

$$\langle \phi_m | H_1 | \phi_n \rangle + (E_m^{(0)} - E_n^{(0)}) C_{nm}^{(1)} = 0$$

that is,

$$\lambda C_{nm}^{(1)} = \frac{\langle \phi_m | \lambda H_1 | \phi_n \rangle}{E_n^{(0)} - E_m^{(0)}} \quad (11-11)$$

The numerator is the matrix element of λH_1 in the basis of states in which H_0 is diagonal. This formula is used in the next equation, which comes from the identification of terms proportional to λ^2 :

$$H_0 \sum_{k \neq n} C_{nk}^{(2)} |\phi_k\rangle + H_1 \sum_{k \neq n} C_{nk}^{(1)} |\phi_k\rangle = E_n^{(0)} \sum_{k \neq n} C_{nk}^{(2)} |\phi_k\rangle + E_n^{(1)} \sum_{k \neq n} C_{nk}^{(1)} |\phi_k\rangle + E_n^{(2)} |\phi_n\rangle \quad (11-12)$$

Taking the scalar product with $\langle \phi_n |$ gives

$$E_n^{(2)} = \sum_{k \neq n} \langle \phi_n | H_1 | \phi_k \rangle C_{nk}^{(1)} = \sum_{k \neq n} \frac{\langle \phi_n | H_1 | \phi_k \rangle \langle \phi_k | H_1 | \phi_n \rangle}{E_n^{(0)} - E_k^{(0)}} = \sum_{k \neq n} \frac{|\langle \phi_n | H_1 | \phi_k \rangle|^2}{E_n^{(0)} - E_k^{(0)}} \quad (11-13)$$

The last step follows from the hermiticity of H_1 :

$$\langle \phi_k | H_1 | \phi_n \rangle = \langle \phi_n | H_1 | \phi_k \rangle^*$$

This, too, is a very important formula, especially since the first-order shift frequently vanishes on grounds of symmetry. We may interpret the formula as follows: The second-order energy shift is the sum of terms, whose strength is given by the square of the matrix element connecting the given state ϕ_n to all other states by the perturbing potential, weighted by the reciprocal of the energy difference between the states. We can draw several conclusions from the formula.

- (a) If $|\phi_n\rangle$ is the *ground state*—that is, the state of lowest energy—then the denominator in the sum is always negative, and hence (11-13) is always negative.
- (b) All other things being equal—that is, if all the matrix elements of H_1 are of roughly the same order of magnitude (which is the kind of guess one would make without more specific knowledge)—then nearby levels have a bigger effect on the second-order energy shift than distant ones have.
- (c) If an important level “ k ”—important in the sense of lying nearby, or of $\langle \phi_k | H_1 | \phi_n \rangle$ being large—lies above the given level “ n ,” then the second-order shift is downward; if it lies below, the shift is upward. We speak of this as a tendency of levels to repel each other.

An expression for $C_{nk}^{(2)}$ may be obtained from (11-12) by taking the scalar product with $\langle \phi_m |$, $m \neq n$, but we shall not require this formula. Also $N(\lambda)$ can be determined from

$$\begin{aligned} \langle \psi_n | \psi_n \rangle &= N^2(\lambda) \left\{ 1 + \lambda^2 \sum_{k \neq n} |C_{nk}^{(1)}|^2 + \dots \right\} \\ &= 1 \end{aligned} \quad (11-14)$$

It is therefore 1 to first order in λ . Hence, to first order in λ , we may write

$$|\psi_n\rangle = |\phi_n\rangle + \sum_{k \neq n} \frac{\langle \phi_k | \lambda H_1 | \phi_n \rangle}{E_n^{(0)} - E_k^{(0)}} |\phi_k\rangle \quad (11-15)$$

a formula that is sometimes useful.

EXAMPLE 11-1

Consider a charged particle in a simple harmonic oscillator, for which $H_0 = p^2/2m + m\omega^2 x^2/2$, subject to a constant electric field so that $H_1 = q\mathcal{E}x$. Calculate the energy shift for the n th level to first and second order in $(q\mathcal{E})$. (*Hint:* Use the operators A and A^+ for the evaluation of the matrix elements.)

SOLUTION

(a) To first order we need to calculate $q\mathcal{E}\langle n|x|n\rangle$. It is easy to show that $\langle n|x|n\rangle = 0$. One way is to use the relation (7-4), from which we get

$$x = \sqrt{\frac{\hbar}{2m\omega}}(A + A^+)$$

and since $A|n\rangle = \sqrt{n}|n-1\rangle$ and $A^+|n\rangle = \sqrt{n+1}|n+1\rangle$ we see that $\langle n|x|n\rangle = 0$. Another way of seeing this is that the matrix element involves the integral

$$\int_{-\infty}^{\infty} \phi_n^*(x)x\phi_n(x)dx = \int_{-\infty}^{\infty} dx x|\phi_n(x)|^2$$

Since $|\phi_n(x)|^2$ is always an even function of x ($\phi_n(-x) = (-1)^n\phi_n(x)$), the integral from $-\infty$ to $+\infty$ involving an *odd* integrand is zero.

(b) The second-order term involves

$$q^2\mathcal{E}^2 \sum_{k \neq n} \frac{|k|x|n\rangle|^2}{\hbar\omega(n-k)} = \frac{q^2\mathcal{E}^2}{\hbar\omega} \frac{\hbar}{2m\omega} \sum_{k \neq n} \frac{|k|A + A^+|n\rangle|^2}{n-k}$$

The only contributions come from $k = n-1$ and $k = n+1$, so that

$$\sum_{k \neq n} \frac{|k|A + A^+|n\rangle|^2}{n-k} = \frac{|\sqrt{n}|^2}{1} + \frac{|\sqrt{n+1}|^2}{-1} = -1$$

and thus

$$E_n^{(2)} = -\frac{q^2\mathcal{E}^2}{2m\omega^2}$$

The result is independent of n . We can check for its correctness by noting that the total potential energy is

$$\frac{1}{2}m\omega^2 x^2 + q\mathcal{E}x = \frac{1}{2}m\omega^2 \left(x^2 + \frac{2q\mathcal{E}}{m\omega^2}x \right) = \frac{1}{2}m\omega^2 \left(x + \frac{q\mathcal{E}}{m\omega^2} \right)^2 - \frac{q^2\mathcal{E}^2}{2m\omega^2}$$

Thus the perturbation shifts the center of the potential by $-q\mathcal{E}/m\omega^2$ and lowers the energy by $q^2\mathcal{E}^2/2m\omega^2$, which agrees with our second-order result.

11-2 DEGENERATE PERTURBATION THEORY

A special problem arises when the unperturbed Hamiltonian has two or more eigenvalues that are the same. Equation (11-11) shows that there is a problem: $C_{nk}^{(1)}$ becomes infinite for those values of n and k for which $E_n^{(0)}$ and $E_k^{(0)}$ are equal. This occurs if some of the eigenvalues of H_0 are *degenerate*. The general procedure is to eliminate the degeneracy first, and then go on as before.

Let us see how to do this for a 3×3 matrix example, where we shall assume that $E_1^{(0)} = E_2^{(0)} (= E^{(0)})$. We have

$$H_0 = \begin{pmatrix} E^{(0)} & 0 & 0 \\ 0 & E^{(0)} & 0 \\ 0 & 0 & E_3^{(0)} \end{pmatrix} \quad (11-16)$$

and

$$\lambda H_1 = \begin{pmatrix} \lambda h_{11} & \lambda h_{12} & \lambda h_{13} \\ \lambda h_{21} & \lambda h_{22} & \lambda h_{23} \\ \lambda h_{31} & \lambda h_{32} & \lambda h_{33} \end{pmatrix} \quad (11-17)$$

subject to the hermiticity requirement that $h_{ij} = h_{ji}^*$.

We can now separate the problem into two parts. First, we diagonalize the upper left corner 2×2 matrix. As noted in Chapter 9, a hermitian matrix H will be diagonalized by a unitary transformation U :

$$U H U^\dagger = H_D$$

A unitary matrix that only affects the 2×2 “box” in the upper left corner will have the form

$$U = \begin{pmatrix} \times & \times & 0 \\ \times & \times & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (11-18)$$

where the $\times$'s denote unknown U_{ij} . The important thing is to note that this U will not affect H_0 : The lower right-hand corner is multiplied by 1. The upper 2×2 left corner of H_0 is of the form $E^{(0)} \times \mathbf{1}$ where $\mathbf{1}$ is the 2×2 unit matrix. As a consequence of $U \mathbf{1} U^\dagger = \mathbf{1}$, that corner is also unchanged. The net effect of the diagonalization is thus to diagonalize the 2×2 matrix

$$\begin{pmatrix} \lambda h_{11} & \lambda h_{12} \\ \lambda h_{21} & \lambda h_{22} \end{pmatrix}$$

The eigenvalues of the 2×2 matrix are easily calculated. We will just state that after the diagonalization, the total matrix takes the form

$$H = \begin{pmatrix} E^{(0)} + \lambda w_1 & 0 & \lambda h_{13} \\ 0 & E^{(0)} + \lambda w_2 & \lambda h_{23} \\ \lambda h_{31} & \lambda h_{32} & E_3^{(0)} \end{pmatrix} \quad (11-19)$$

with

$$w_1, w_2 = \frac{h_{11} + h_{22}}{2} \pm \sqrt{\left(\frac{h_{11} - h_{22}}{2}\right)^2 + h_{12}h_{21}} \quad (11-20)$$

We have removed the degeneracy, and now the formulas derived above apply without any problems.

If we had started with a 5×5 matrix, and three eigenvalues of H_0 were equal, then we would have to diagonalize a 3×3 submatrix of λH_1 . This would be much messier, but not different in principle. Similarly, if in our 3×3 example all three eigenvalues of H_0 were equal, we would have to diagonalize the 3×3 matrix λH_1 and we would then get an *exact* result, since H would then be diagonal.

Let us discuss degenerate perturbation theory further. The problem is always that instead of a unique $|\phi_n\rangle$ there is now a finite set of $|\phi_n^{(i)}\rangle$, all of which have the same energy $E_n^{(0)}$. This set can be made orthonormal with respect to the label “ i ,” because, as we saw in Section 5-4, this label can be associated with the eigenvalues of some other, simultaneously commuting hermitian operator. We thus choose the set of $|\phi_n^{(i)}\rangle$ such that

$$\langle \phi_m^{(i)} | \phi_n^{(j)} \rangle = \delta_{mn} \delta_{ij} \quad (11-21)$$

The natural way to take the degeneracy into account is to replace (11-3) by an expression that involves linear combinations of the degenerate eigenfunctions of H_0 :

$$|\psi_n\rangle = N(\lambda) \left\{ \sum_i \alpha_i |\phi_n^{(i)}\rangle + \lambda \sum_{k \neq n} C_{nk}^{(1)} \sum_i \beta_i |\phi_k^{(i)}\rangle + \dots \right\} \quad (11-22)$$

The coefficients $\alpha_1, \beta_1, \dots$ will have to be determined. When (11-22) is inserted into the Schrödinger equation (11-2), we get, to first order in λ ,

$$H_0 \sum_{k \neq n} C_{nk}^{(1)} \sum_i \beta_i |\phi_k^{(i)}\rangle + H_1 \sum_i \alpha_i |\phi_n^{(i)}\rangle = E_n^{(1)} \sum_i \alpha_i |\phi_n^{(i)}\rangle + E_n^{(0)} \sum_{k \neq n} C_{nk}^{(1)} \sum_i \beta_i |\phi_k^{(i)}\rangle \quad (11-23)$$

Taking the scalar product with $\langle \phi_n^{(j)} |$ gives the first-order shift equation,

$$\sum_i \langle \phi_n^{(j)} | \lambda H_1 | \phi_n^{(i)} \rangle = \lambda E_n^{(1)} \alpha_j \quad (11-24)$$

This is a finite-dimensional eigenvalue problem. For example, if there is a twofold degeneracy, and if we use the notation $\langle \phi_n^{(i)} | H_1 | \phi_n^{(j)} \rangle = h_{ji}$, the equation reads

$$\begin{aligned} h_{11}\alpha_1 + h_{12}\alpha_2 &= E_n^{(1)}\alpha_1 \\ h_{21}\alpha_1 + h_{22}\alpha_2 &= E_n^{(1)}\alpha_2 \end{aligned}$$

The eigenvalues and the α_i can be determined from this equation, if we add the condition that

$$\sum_i |\alpha_i|^2 = 1$$

We do not bother with the determination of the β_i , since we shall only use degenerate perturbation theory for the first-order energy eigenvalues in our applications. If it so happens that $h_{ij} = 0$ for $i \neq j$ —that is, the matrix h_{ij} is diagonal—then the first-order shifts are just the diagonal elements of this matrix. This will happen when the perturbation H_1 commutes with the operator whose eigenvalues the “ i ” labels represent. For example, in the hydrogen atom, there is a degeneracy associated with the eigenvalues of L_z ; that is, all m -values have the same energy. If it happens that

$$[H_1, L_z] = 0 \quad (11-25)$$

and if we choose our $\phi_n^{(i)}$ to be eigenfunctions of L_z , then h_{ij} will be diagonal. To see this, note that with

$$\begin{aligned} L_z |\phi_n^{(i)}\rangle &= \hbar m^{(i)} |\phi_n^{(i)}\rangle \\ \langle \phi_n^{(j)} | [H_1, L_z] | \phi_n^{(i)} \rangle &= \langle \phi_n^{(j)} | H_1 L_z - L_z H_1 | \phi_n^{(i)} \rangle \\ &= \hbar (m^{(i)} - m^{(j)}) h_{ji} \\ &= 0 \end{aligned} \quad (11-26)$$

that is, (11-26) implies

$$h_{ij} = 0 \quad \text{for } m^{(i)} \neq m^{(j)} \quad (11-27)$$

An example of such a situation occurs when a rigid rotator described by the Hamiltonian

$$H_0 = \frac{\mathbf{L}^2}{2I} \quad (11-28)$$

with magnetic moment

$$\boldsymbol{\mu} = \frac{qg}{2M} \mathbf{L}$$

is placed in a magnetic field that points in the z -direction. The perturbation is of the form

$$H_1 = -\frac{qg}{2M} BL_z \quad (11-29)$$

Because the eigenfunctions of H_0 are the Y_{lm} , and these are simultaneously eigenfunctions of L_z , we can immediately read off the eigenvalues. These are

$$E = \frac{l(l+1)\hbar^2}{2I} - \frac{qgB\hbar}{2M} m_l \quad -l \leq m_l \leq l \quad (11-30)$$

A more detailed discussion of degenerate perturbation theory appears in our treatment of the Stark effect. Other illustrations will appear in our discussion of the real hydrogen atom in Chapter 12.

11-3 THE STARK EFFECT

To illustrate the application of perturbation theory to a real problem, we will consider the effect of an external electric field on the energy levels of a hydrogenlike atom. This is the *Stark effect*. The unperturbed Hamiltonian is

$$H_0 = \frac{\mathbf{p}^2}{2\mu} - \frac{Ze^2}{4\pi\epsilon_0 r} \quad (11-31)$$

whose eigenfunctions we denote by $\phi_{nlm}(\mathbf{r})$. The perturbing potential is

$$\lambda H_1 = e\mathcal{E} \cdot \mathbf{r} = e\mathcal{E}z \quad (11-32)$$

where $\mathcal{E}$ is the electric field. The quantity $e\mathcal{E}$ will play the role of the parameter λ . The energy shift of the ground state, which is nondegenerate, is given by the expression

$$E_{100}^{(1)} = e\mathcal{E}\langle\phi_{100}|z|\phi_{100}\rangle = e\mathcal{E} \int d^3r |\phi_{100}(r)|^2 z \quad (11-33)$$

This integral vanishes, since the square of the wave function is always an even function under parity, and the perturbing potential is an odd function under reflections. Thus for the ground state there is no energy shift that is linear in the electric field $\mathcal{E}$. Classically, a system that has an electric dipole moment, $\mathbf{d}$, will experience an energy shift of magnitude $-\mathbf{d} \cdot \mathcal{E}$. Thus (11-33) shows that in its ground state the atom has no permanent dipole moment. The parity argument can be used whenever the unperturbed Hamiltonian is invariant under reflection, and it can be generalized to the statement that *systems in nondegenerate states cannot have permanent dipole moments*. The statement of nondegeneracy is important: It is only then that the states are also eigenstates of the parity operator, and then $|\phi(\mathbf{r})|^2$ is even, and the expectation value of z vanishes.

There are molecules that have permanent electric dipole moments. The explanation for this is that these molecules have degenerate ground states with opposite parity. If these states are described by $|\psi_{\pm}\rangle$, the expectation value of z in a state like $\alpha|\psi_{+}\rangle + \beta|\psi_{-}\rangle$ will not vanish:

$$\langle \alpha\psi_{+} + \beta\psi_{-} | z | \alpha\psi_{+} + \beta\psi_{-} \rangle = \alpha^{*}\beta\langle \psi_{+} | z | \psi_{-} \rangle + \alpha\beta^{*}\langle \psi_{-} | z | \psi_{+} \rangle \quad (11-34)$$

What happens if the states ψ_{+} and ψ_{-} have energies that are split by a very small amount? This, after all, is true in most cases where the states are degenerate only in first approximation. The degeneracy may be removed by perturbations that affect the two states differently, or by tunneling, as was discussed in Chapter 4. We will see below [see eq. (11-53)] that for electric fields strong enough that $e\langle \mathbf{r} \rangle \cdot \mathbf{E}$ is large compared with the energy by which the states are split, then the change in energy will be linear in E , the electric field, and the system may be described by an electric dipole moment.

Let us look at the second-order term. It reads

$$E_{100} = e^2 \mathcal{E}^2 \left\{ \sum_{nlm} \frac{|\langle \phi_{nlm} | z | \phi_{100} \rangle|^2}{E_1^{(0)} - E_n^{(0)}} + \sum_k \frac{|\langle \phi_k | z | \phi_{100} \rangle|^2}{E_1^{(0)} - \hbar^2 k^2 / 2m} \right\} \quad (11-35)$$

The reason for the second term is that in (11-35) we must sum over a complete set of eigenstates of H_0 . For atoms, these consist of the bound states $|\phi_{nlm}\rangle$, as well as the continuum states in which the electron has a positive energy. We label the continuum states by $\mathbf{k}$, where k is related to the positive kinetic energy by $E = \hbar^2 k^2 / 2m$. This sum is very difficult to evaluate directly, since it involves an integral over k , which appears in the rather complicated continuum solutions of the Coulomb problem (it can be done by a trick, which we shall not describe). What we can do, however, is to estimate the value of $E_{100}^{(2)}$ by finding an upper bound. Let us rewrite (11-35) in a more symbolic form

$$E_{100}^{(2)} = e^2 \mathcal{E}^2 \sum_{E \neq E_1} \frac{|\langle \phi_{100} | z | \phi_E \rangle|^2}{E_1^{(0)} - E} \quad (11-36)$$

where the complete set is now labeled by $|\phi_E\rangle$. Notice that because $E_1^{(0)}$ is the lowest energy $E_{100}^{(2)}$ is negative. Because all the energies E lie above $E_1^{(0)}$, we have

$$E - E_1^{(0)} \geq E_2^{(0)} - E_1^{(0)}$$

so that

$$\frac{1}{E - E_1^{(0)}} \leq \frac{1}{E_2^{(0)} - E_1^{(0)}}$$

Consequently

$$-E_{100}^{(2)} \leq e^2 \mathcal{E}^2 \frac{1}{E_2^{(0)} - E_1^{(0)}} \sum_E \langle \phi_{100} | z | \phi_E \rangle \langle \phi_E | z | \phi_{100} \rangle \quad (11-37)$$

We can include the ground state in the sum since $\langle \phi_{100} | z | \phi_{100} \rangle = 0$. The sum can now be simplified with the help of the completeness relation $\sum_E |\phi_E\rangle \langle \phi_E| = \mathbf{1}$, so that

$$\sum_E \langle \phi_{100} | z | \phi_E \rangle \langle \phi_E | z | \phi_{100} \rangle = \langle \phi_{100} | z^2 | \phi_{100} \rangle \quad (11-38)$$

This is easily evaluated. Because the ground-state wave function is spherically symmetric

$$\begin{aligned} \langle \phi_{100} | z^2 | \phi_{100} \rangle &= \langle \phi_{100} | x^2 | \phi_{100} \rangle = \langle \phi_{100} | y^2 | \phi_{100} \rangle \\ &= \frac{1}{3} \langle \phi_{100} | x^2 + y^2 + z^2 | \phi_{100} \rangle = \frac{1}{3} \langle \phi_{100} | r^2 | \phi_{100} \rangle = a_0^2 \end{aligned} \quad (11-39)$$

Since

$$E_2^{(0)} - E_1^{(0)} = -\frac{1}{2} mc^2 \alpha^2 \left(\frac{1}{4} - 1 \right) = \frac{8}{3} mc^2 \alpha^2$$

the inequality becomes

$$-E_{100}^{(2)} < \frac{8e^2 \mathcal{E}^2 a_0^2}{3mc^2 \alpha^2} = \frac{8}{3} (4\pi \epsilon_0 \mathcal{E}^2) a_0^3 \quad (11-40)$$

Note that $4\pi \epsilon_0 \mathcal{E}^2 \times \text{volume}$ has the dimensions of an energy. Since the Bohr radius a_0 is the only length in the problem, we expect from dimensional analysis that

$$-E_{100}^{(2)} = (\text{const})(4\pi \epsilon_0 \mathcal{E}^2 a_0^3) \quad (11-41)$$

The exact second-order result gives the value 9/4 for the constant.

The Stark Effect for $n = 2$ States

To illustrate degenerate perturbation theory, we calculate the first-order (linear in $\mathcal{E}$) Stark effect for the $n = 2$ states of the hydrogen atom. For the unperturbed system there are really four $n = 2$ states that have the same energy. These are

$$\begin{aligned} \phi_{200} &= (2a_0)^{-3/2} 2 \left(1 - \frac{r}{2a_0} \right) e^{-r/2a_0} Y_{00} \\ \phi_{211} &= (2a_0)^{-3/2} 3^{-1/2} \left(\frac{r}{a_0} \right) e^{-r/2a_0} Y_{11} \\ \phi_{210} &= (2a_0)^{-3/2} 3^{-1/2} \left(\frac{r}{a_0} \right) e^{-r/2a_0} Y_{10} \\ \phi_{2,1,-1} &= (2a_0)^{-3/2} 3^{-1/2} \left(\frac{r}{a_0} \right) e^{-r/2a_0} Y_{1,-1} \end{aligned} \quad (11-42)$$

The $l = 0$ state has even parity, and the $l = 1$ state has odd parity. We want to solve an equation like (11-24) and, on the face of it, four equations are involved. If we note, however, that (1) the perturbing potential (that is, z) commutes with L_z so that it only connects states with the same m -value, and (2) parity forces us to consider only terms in which the perturbing potential must connect $l = 1$ to $l = 0$ terms—that is,

$$\langle \phi_{2,1,\pm 1} | z | \phi_{2,1,\pm 1} \rangle = 0$$

then the matrix in (11-24) is only a 2×2 matrix. The equation reads

$$e\mathcal{E} \begin{pmatrix} \langle \phi_{200} | z | \phi_{200} \rangle & \langle \phi_{200} | z | \phi_{210} \rangle \\ \langle \phi_{210} | z | \phi_{200} \rangle & \langle \phi_{210} | z | \phi_{210} \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = E^{(1)} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad (11-43)$$

The diagonal elements are zero, because of parity, and the off-diagonal elements are equal, since they are complex conjugates of each other, and each may be chosen to be real. We have

$$\begin{aligned} \langle \phi_{200} | z | \phi_{210} \rangle &= \int_0^\infty r^2 dr (2a_0)^{-3} e^{-r/a_0} \frac{2r}{\sqrt{3}a_0} \left(1 - \frac{r}{2a_0} \right) r \\ &\quad \cdot \int d\Omega Y_{00}^* (\sqrt{4\pi/3} Y_{10}) Y_{10} \\ &= -3a_0 \end{aligned} \quad (11-44)$$

and hence (11-43) becomes

$$\begin{pmatrix} -E^{(1)} & -3e\mathcal{E}a_0 \\ -3e\mathcal{E}a_0 & -E^{(1)} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = 0 \quad (11-45)$$

The eigenvalues of this are

$$E^{(1)} = \pm 3e\mathcal{E}a_0 \quad (11-46)$$

and the corresponding eigenstates, when properly normalized, are

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ and } \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (11-47)$$

respectively. Thus the linear Stark effect for the $n = 2$ states yields a splitting of degenerate levels as shown in Fig. 11-1.

There are some general comments that can be abstracted from the calculations just concluded.

- (a) The states in the presence of the electric field are no longer eigenstates of $\mathbf{L}^2$. In the preceding case, for example, we found that the states that diagonalize the perturbation were equal mixtures of $l = 0$ and $l = 1$, though they are still eigenstates of L_z . The reason is that the perturbation changes the Hamiltonian, so that it no longer commutes with $\mathbf{L}^2$. This can be worked out in detail, but it is really evident that the external field specifies a preferred direction, so that the physical system is no longer invariant under arbitrary rotations. It is still invariant under rotations about the preferred axis, here the z -direction, and hence L_z is still a good constant of the motion.
- (b) Quite generally, whenever there is a perturbation that does not conserve some quantity (for example, $\mathbf{L}^2$ here), then the states that “diagonalize” the new Hamiltonian in any approximation, are superpositions of states with different values of the previously conserved quantum numbers, and thus degenerate levels will be split.
- (c) We can summarize the procedure in degenerate perturbation theory in matrix language as follows. If H_0 is diagonal, but H_1 is not, then, since H_0 and H_1 do not commute, it is not possible to diagonalize H_1 by itself, without “undagonalizing” H_0 . One must work with

$$H = H_0 + H_1$$

as a whole. If we work with a subset of degenerate states, all of which are eigenstates of H_0 with the same eigenvalue, then, as far as these states are concerned, H_0 is not merely diagonal, but it is proportional to the unit matrix. Since H_1 (and everything else) commutes with the unit matrix, one can diagonalize H_1 by itself, without affecting H_0 .

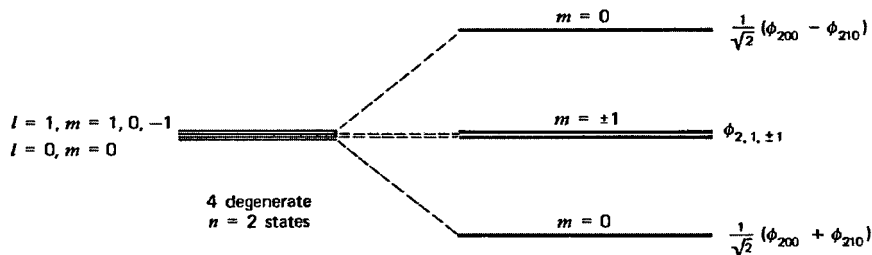


Figure 11-1 Pattern of Stark splitting of hydrogen atom in $n = 2$ state. The fourfold degeneracy is partly lifted by the perturbation. The $m = \pm 1$ states remain degenerate and are not shifted in the Stark effect.

The hydrogenlike atoms considered here were somewhat idealized. As we will see in Chapter 12, there are small relativistic and spin-orbit coupling effects that actually remove some of the degeneracies. Does this mean that we never really need to use degenerate perturbation theory? Actually, even if, say, ϕ_{200} and ϕ_{210} do not have exactly the same energy, it may still be sensible to take some linear combination of them in the perturbation expansion. If we have, for example

$$\begin{aligned} H_0 \phi_{200} &= (E_2^0 - \Delta) \phi_{200} \\ H_0 \phi_{210} &= (E_2^0 + \Delta) \phi_{210} \end{aligned} \quad (11-48)$$

with Δ small, then the Schrödinger equation, with the linear combinations, reads

$$(H_0 + \lambda H_1) \left(\alpha_1 \phi_{200} + \alpha_2 \phi_{210} + \lambda \sum_{n \neq 2} C_n \phi_n \right) = E \left(\alpha_1 \phi_{200} + \alpha_2 \phi_{210} + \lambda \sum_{n \neq 2} C_n \phi_n \right) \quad (11-49)$$

Taking the scalar product with ϕ_{200} and ϕ_{210} , respectively, leads to the following equation to order λ :

$$\begin{pmatrix} E_2^0 - \Delta + \langle \phi_{200} | \lambda H_1 | \phi_{200} \rangle & \langle \phi_{200} | \lambda H_1 | \phi_{210} \rangle \\ \langle \phi_{210} | \lambda H_1 | \phi_{200} \rangle & E_2^0 + \Delta + \langle \phi_{210} | \lambda H_1 | \phi_{210} \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = E \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad (11-50)$$

If we write

$$\langle \phi_{200} | \lambda H_1 | \phi_{210} \rangle = \langle \phi_{210} | \lambda H_1 | \phi_{200} \rangle = a\lambda \quad (11-51)$$

we must find the eigenvalues of the matrix

$$\begin{pmatrix} E_2^0 - \Delta & \lambda a \\ \lambda a & E_2^0 + \Delta \end{pmatrix} \quad (11-52)$$

and these are

$$E = E_2^0 \pm \sqrt{a^2 \lambda^2 + \Delta^2} \quad (11-53)$$

(In the preceding we have used $\langle \phi_{200} | H_1 | \phi_{200} \rangle = \langle \phi_{210} | H_1 | \phi_{210} \rangle = 0$.) We see that when $\Delta \gg a\lambda$, we get a “quadratic” effect only. This corresponds to no degeneracy. When $\Delta \ll a\lambda$ we get the result of the form (11-46). In the intermediate region, the preceding, more careful treatment is necessary. Furthermore, when the new linear combinations are used, then in second-order perturbation theory there no longer appear very tiny energy differences in the denominators. We do not discuss this in detail, but this is not difficult to establish.

Some Final Comments

Let us consider the shift of the ground state. The second-order perturbation term involves $4\pi\epsilon_0\mathcal{E}^2 a_0^3$, which has the dimensions of an energy. Higher order terms must involve higher powers of $\mathcal{E}^2$ and still have the dimensions of an energy. This means that the perturbation series for the shift must be of the form

$$\Delta E_{100} = 4\pi\epsilon_0\mathcal{E}^2 a_0^3 (A_0 + A_1\beta + A_2\beta^2 + \cdots) \quad (11-54)$$

where

$$\beta = \frac{4\pi\epsilon_0\mathcal{E}^2 a_0^3}{mc^2\alpha^2} \quad (11-55)$$

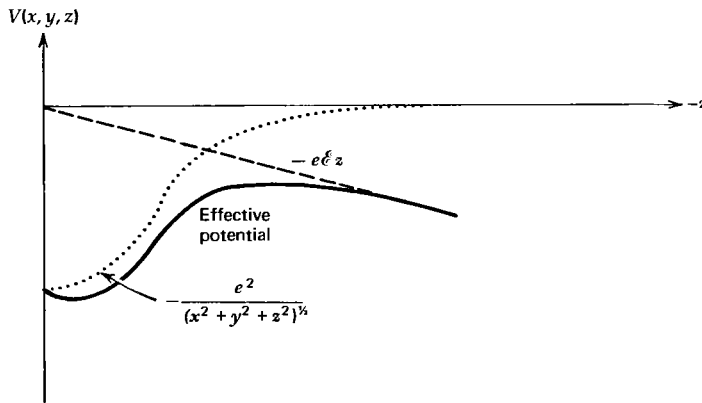


Figure 11-2 Schematic picture of potential energy as a function of z with x and y held fixed. The dotted line represents the Coulomb potential, the dashed line the potential energy due to the external field, and the solid line the total potential.

is a dimensionless parameter involving the ratio of the only two energies that enter into the problem. The first two terms are actually in excellent agreement with experiment. Can we expect the series to converge? The answer is that it does not converge. The reason can be seen in Fig. 11-2, which is a slice of the potential energy for x and y held fixed. It is evident that there is a barrier rather than just a well for the bound electron. The barrier is ultimately penetrable, even though for small β it is very broad. Once an electron tunnels through the barrier, the external field accelerates it, so that in nonrelativistic quantum mechanics, at least, it can acquire an arbitrarily large kinetic energy. Since the tunneling time is large, even on the scale of the age of the universe, the levels are *metastable*. The mathematical description of the series in (11-54) is that it is an *asymptotic series*, in which the first few terms give an excellent description of the real energy shifts.

PROBLEMS

1. Consider a simple harmonic oscillator in one dimension with characteristic frequency ω . Suppose a perturbation λx^2 is introduced. Calculate the energy shift of the n th level to first order in λ . Can you work out the second-order term, proportional to λ^2 , without actually doing the calculation? (*Hint*: Write out the *total* Hamiltonian, including the perturbation.)
2. Consider a symmetric rotator with $H_0 = L^2/2I$. Suppose this system is subject to a perturbation given by

$$H_1 = E_1 \cos \theta$$

What are the energy shifts for the states with $l = 1$?

3. Consider a particle in an infinite potential well with width L with one edge at $x = 0$. What is the energy shift due to the introduction of an additional potential, which corresponds to tilting the "floor" of the well,

$$V(x) = V_0(x/L)$$

for the n th state?

4. Diagonalize the matrix given by

$$\begin{pmatrix} E & \lambda & 0 & 0 \\ \lambda & E & 0 & 0 \\ 0 & 0 & 2E & \sigma \\ 0 & 0 & \sigma & 0 \end{pmatrix}$$

5. Consider the hydrogen atom and assume that the proton, instead of being as point-source of the Coulomb field, is a uniformly charged sphere of radius R . This means that the Coulomb potential is now modified so that

$$V(r) = -\frac{3e^2}{8\pi\epsilon_0 R^3} \left(R^2 - \frac{1}{3} r^2 \right) \quad r < R \quad (\ll a_0)$$

$$= -\frac{e^2}{4\pi\epsilon_0 r} \quad r > R$$

Calculate the energy shift for the $n = 1, l = 0$ state, and for the $n = 2$ states caused by this modification, using the wave functions given in Chapter 8 (p. 138).

6. Calculate the energy shift in the ground state of the one-dimensional harmonic oscillator, when the perturbation

$$V = \lambda x^4$$

is added to

$$H = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 x^2$$

7. The bottom of an infinite well is changed to have the shape

$$V(x) = \epsilon \sin \frac{\pi x}{b} \quad 0 \leq x \leq b$$

Calculate the energy shifts for all the excited states to first order in ϵ . Note that the well originally had $V(x) = 0$ for $0 \leq x \leq b$, with $V = \infty$ elsewhere.

8. Prove the sum rule (Thomas-Reiche-Kuhn sum rule)

$$\sum_n (E_n - E_a) |\langle n | x | a \rangle|^2 = \frac{\hbar^2}{2m}$$

[Hints: (a) Write the commutation relation $[p, x] = \hbar/i$ in the form

$$\sum_n [\langle a | p | n \rangle \langle n | x | a \rangle - \langle a | x | n \rangle \langle n | p | a \rangle] = \frac{\hbar}{i} \langle a | a \rangle = \frac{\hbar}{i}$$

(b) Use the fact that

$$\langle a | p | n \rangle = \left\langle a \left| m \frac{dx}{dt} \right| n \right\rangle = m \frac{i}{\hbar} \langle a | [H, x] | n \rangle$$

in working out the problem.]

9. Check the sum rule in Problem 8 for the one-dimensional harmonic oscillator, with “ a ” taken in the ground state.
10. Work out the first-order Stark effect in the $n = 3$ state of the hydrogen atom. Do not bother to work out all the integrals, but you should construct the correct linear combinations of states. Can you give a qualitative explanation of the pattern of energy shifts?
11. Consider a two-dimensional harmonic oscillator described by the Hamiltonian

$$H = \frac{1}{2m} (p_x^2 + p_y^2) + \frac{1}{2} m \omega^2 (x^2 + y^2)$$

Generalize the approach of Chapter 6 to obtain solutions of this problem in terms of raising operators acting on the ground state. Calculate the energy shifts due to the perturbation

$$V = 2\lambda xy$$

in the ground state, and in the degenerate first excited states, using first-order perturbation theory. Can you interpret your result very simply? Solve the problem exactly, and compare it with a second-order perturbation calculation of the shift of the ground state.

12. The Hamiltonian for an electron in a hydrogen atom subject to a constant magnetic field $\mathbf{B}$ is, with the neglect of spin,

$$H = \frac{\mathbf{p}^2}{2m_e} - \frac{e^2}{4\pi\epsilon_0 r} + \frac{e}{2m} \mathbf{L} \cdot \mathbf{B}$$

where $\mathbf{L}$ is the angular momentum operator. In the absence of the magnetic field, there will be a single line in the transition from an $(n = 4, l = 3)$ state to an $(n = 3, l = 2)$ state. What will be the effect of the magnetic field on that line? Sketch the new spectrum and the possible transitions, constrained by the selection rules $\Delta l_z = 0, \pm 1$. How many lines will there be? What will be the effect of a constant electric field $\mathbf{E}$ parallel to $\mathbf{B}$?

13. Consider a Hamiltonian of the form

$$H = \begin{pmatrix} E_0 & 0 \\ 0 & -E_0 \end{pmatrix} + \lambda \begin{pmatrix} \alpha & u \\ u^* & \beta \end{pmatrix}$$

- (a) Calculate the energy shift to first and second order in λ . Compare your results with the exact eigenvalues.
- (b) Suppose u^* is replaced by $v \neq u^*$. Show that the eigenstates of the non-hermitian new H corresponding to different eigenvalues are no longer orthogonal. (For this part of the problem you may take $\alpha = \beta = 0$ to simplify the work.)

Chapter 12

The Real Hydrogen Atom

The discussion of hydrogenlike atoms in Chapter 8 was based on the Hamiltonian

$$H_0 = \frac{\mathbf{p}^2}{2\mu} - \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{r} \quad (12-1)$$

In a more realistic treatment, several corrections must be taken into account. We separate out the treatment of the proton motion so as not to clutter up our discussion of relativistic and spin effects with simple kinematic matters that involve the motion of the proton. For the time being we treat the proton as infinitely massive, and return to the effect of its motion in a separate section.

12-1 RELATIVISTIC KINETIC ENERGY EFFECTS

The relativistic expression for the kinetic energy of the electron is

$$K = \sqrt{(\mathbf{p}c)^2 + (m_e c^2)^2} - m_e c^2 \approx \frac{\mathbf{p}^2}{2m_e} - \frac{1}{8} \frac{(\mathbf{p}^2)^2}{m_e^3 c^2} + \dots \quad (12-2)$$

The second term is the perturbation we want to consider. It gives rise to

$$H_1 = -\frac{1}{2} \left(\frac{\mathbf{p}^2}{2m_e} \right)^2 \frac{1}{m_e c^2} \quad (12-3)$$

We can estimate its effect on the basic energy levels. We know that $\mathbf{p}^2/2m_e \approx O(\alpha^2 m_e c^2)$ so that the perturbation is of order α^2 smaller than the leading term. We will evaluate it more accurately a little later.

12-2 SPIN-ORBIT COUPLING

The spin of the electron and its associated magnetic moment change the interaction of the electron with the electric field due to the proton charge. An electron moving with uniform velocity $\mathbf{v}$ experiences a magnetic field

$$\mathbf{B} = -\frac{\mathbf{v} \times \mathbf{E}}{c^2} = -\frac{\mathbf{v} \times (-\nabla\phi(r))}{c^2} = \frac{1}{c^2} \mathbf{v} \times \frac{\mathbf{r}}{r} \frac{d\phi(r)}{dr} \quad (12-4)$$

where $\phi(r)$ is the potential due to the proton charge. The magnetic moment of the electron interacts with this magnetic field. With

$$\boldsymbol{\mu} = -\frac{e}{2m_e} g\mathbf{S} = -\frac{e}{m_e} \mathbf{S}$$

the additional interaction is

$$\begin{aligned} -\boldsymbol{\mu} \cdot \mathbf{B} &= \frac{e}{m_e} \mathbf{S} \cdot \mathbf{v} \times \mathbf{r} \frac{1}{c^2} \frac{1}{r} \frac{d\phi(r)}{dr} = \frac{e}{m_e^2 c^2} \mathbf{S} \cdot \mathbf{p} \times \mathbf{r} \frac{1}{r} \frac{d\phi(r)}{dr} \\ &= -\frac{e}{m_e^2 c^2} \mathbf{S} \cdot \mathbf{L} \frac{1}{r} \frac{d\phi(r)}{dr} \end{aligned} \quad (12-5)$$

For a potential due to a point charge, $\phi(r) = Ze/4\pi\epsilon_0 r$, this becomes

$$\frac{Ze^2}{4\pi\epsilon_0} \frac{1}{m_e^2 c^2} \frac{\mathbf{S} \cdot \mathbf{L}}{r^3} \quad (12-6)$$

Actually this is not correct. It turns out that relativistic effects associated with the fact that the electron does not move in a straight line reduce the above by a factor of 2.¹ This means that the correct perturbation is

$$H_2 = \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{2m_e^2 c^2} \frac{\mathbf{S} \cdot \mathbf{L}}{r^3} \quad (12-7)$$

Let us now use first-order perturbation theory to calculate the effects of H_1 and H_2 on the spectrum of hydrogenlike atoms. We can rewrite H_1 in the form

$$H_1 = -\frac{1}{8} \frac{(\mathbf{p}^2)^2}{m_e^3 c^2} = -\frac{1}{2m_e c^2} \left(\frac{\mathbf{p}^2}{2m_e} \right)^2 = -\frac{1}{2m_e c^2} \left(H_0 + \frac{Ze^2}{4\pi\epsilon_0 r} \right) \left(H_0 + \frac{Ze^2}{4\pi\epsilon_0 r} \right) \quad (12-8)$$

if we neglect reduced mass effects in H_1 . We can now evaluate

$$\begin{aligned} \langle \phi_{nlm} | H_1 | \phi_{nlm} \rangle &= -\frac{1}{2m_e c^2} \left\langle \phi_{nlm} \left| \left(H_0 + \frac{Ze^2}{4\pi\epsilon_0 r} \right) \left(H_0 + \frac{Ze^2}{4\pi\epsilon_0 r} \right) \right| \phi_{nlm} \right\rangle \\ &= -\frac{1}{2m_e c^2} \left[E_n^2 + 2E_n \left\langle \phi_{nlm} \left| \frac{Ze^2}{4\pi\epsilon_0 r} \right| \phi_{nlm} \right\rangle + \left\langle \phi_{nlm} \left| \left(\frac{Ze^2}{4\pi\epsilon_0 r} \right)^2 \right| \phi_{nlm} \right\rangle \right] \\ &= -\frac{1}{2m_e c^2} \left\{ \left[\frac{m_e c^2 (Z\alpha)^2}{2n^2} \right]^2 - 2 \left[\frac{m_e c^2 (Z\alpha)^2}{2n^2} \right] \frac{Ze^2}{4\pi\epsilon_0} \left(\frac{Z}{a_0 n^2} \right) \right. \\ &\quad \left. + \left(\frac{Ze^2}{4\pi\epsilon_0} \right)^2 \frac{2Z^2}{a_0^2 n^3 (2l+1)} \right\} \\ &= -\frac{1}{2} m_e c^2 (Z\alpha)^2 \left[\frac{2(Z\alpha)^2}{n^3 (2l+1)} - \frac{3(Z\alpha)^2}{4n^4} \right] \end{aligned} \quad (12-9)$$

In calculating the preceding we have used the expressions for $\langle \phi_{nlm} | 1/r | \phi_{nlm} \rangle$ and $\langle \phi_{nlm} | 1/r^2 | \phi_{nlm} \rangle$ obtained in (8-52). The spin of the electron does not enter into this part of the energy shift, since H_1 does not depend on the spin. H_2 does depend on the spin, and for our unperturbed wave functions we must take the two-component wave functions in order to calculate the expectation value of H_2 .

¹The so-called *Thomas precession* factor, a relativistic effect calculated by L. H. Thomas, led to the correct doublet splitting in spectra and it convinced the experts of the correctness of the notion of spin. A classical discussion may be found in J. D. Jackson, *Classical Electrodynamics*, Wiley, New York, 1998. The spin-orbit term emerges naturally from the relativistic Dirac equation.

Here, again, we have an example of degenerate perturbation theory. For a given n and l , there are $2(2l + 1)$ degenerate eigenstates of H_0 , with the additional factor of 2 coming from the two spin states. Thus the calculation of the energy shift involves a diagonalization of a submatrix, as in (11-24). We can save ourselves a great deal of labor by noting that

$$\mathbf{S} + \mathbf{L} = \mathbf{J} \quad (12-10)$$

implies that

$$\mathbf{S}^2 + 2\mathbf{S} \cdot \mathbf{L} + \mathbf{L}^2 = \mathbf{J}^2$$

that is,

$$\mathbf{S} \cdot \mathbf{L} = \frac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \quad (12-11)$$

Thus if we combine the degenerate eigenfunctions into linear combinations that are eigenfunctions of $\mathbf{J}^2$ (they already are eigenfunctions of $J_z = L_z + S_z$), then these linear combinations will diagonalize H_2 . The appropriate linear combinations were obtained in Supplement 10-A [Eqs. (10A-13) and (10A-14)]. [www.wiley.com/college/gasiorowicz]

With these linear combinations, which we label by their j and m_j values in terms of l , we have

$$\begin{aligned} \mathbf{S} \cdot \mathbf{L} \left| l + \frac{1}{2}, m + \frac{1}{2} \right\rangle &= \frac{1}{2} (\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \left| l + \frac{1}{2}, m + \frac{1}{2} \right\rangle \\ &= \frac{1}{2} \hbar^2 \left[\left(l + \frac{1}{2} \right) \left(l + \frac{3}{2} \right) - l(l+1) - \frac{1}{2} \frac{3}{2} \right] \left| l + \frac{1}{2}, m + \frac{1}{2} \right\rangle \quad (12-12) \\ &= \frac{1}{2} \hbar^2 l \left| l + \frac{1}{2}, m + \frac{1}{2} \right\rangle \end{aligned}$$

and

$$\begin{aligned} \mathbf{S} \cdot \mathbf{L} \left| l - \frac{1}{2}, m + \frac{1}{2} \right\rangle &= \frac{1}{2} (\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \left| l - \frac{1}{2}, m + \frac{1}{2} \right\rangle \\ &= \frac{1}{2} \hbar^2 \left[\left(l - \frac{1}{2} \right) \left(l + \frac{1}{2} \right) - l(l+1) - \frac{1}{2} \frac{3}{2} \right] \left| l - \frac{1}{2}, m + \frac{1}{2} \right\rangle \quad (12-13) \\ &= -\frac{1}{2} \hbar^2 (l+1) \left| l - \frac{1}{2}, m + \frac{1}{2} \right\rangle \end{aligned}$$

For a given l value there are $[2(l + 1/2) + 1]$ and $[2(l - 1/2) + 1]$ states. What has happened is that the degenerate states have merely been rearranged, but the two groups that they have been split into behave differently under the action of H_2 . If we call the linear combinations $|\phi_{jm,l}\rangle$, then

$$\langle \phi_{jm,l} | H_2 | \phi_{jm,l} \rangle = \frac{1}{2m_e^2 c^2} \frac{Ze^2}{4\pi\epsilon_0} \frac{\hbar^2}{2} \left\{ \begin{matrix} l \\ -l-1 \end{matrix} \right\} \int_0^\infty r^2 dr (R_{nl}(r))^2 \frac{1}{r^3} \quad (12-14)$$

for $j = l \pm 1/2$, respectively. We look up the integral in (8-52) and find that

$$\left\langle \frac{1}{r^3} \right\rangle_{nl} = \frac{Z^3}{a_0^3} \frac{2}{n^3 l(l+1)(2l+1)} \quad (12-15)$$

which is valid for $l \neq 0$. Note that here $a_0 = \hbar/m_e c \alpha$.

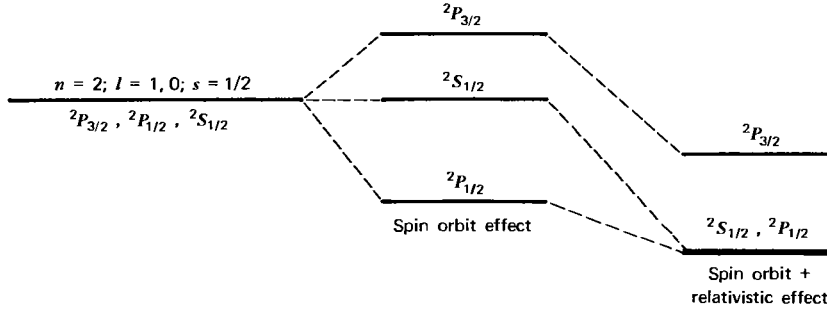


Figure 12-1 Splitting of the $n = 2$ levels by (1) the spin-orbit coupling (which leaves the S state unaffected) and (2) the relativistic effect. The final degeneracy of the $^2S_{1/2}$ and $^2P_{1/2}$ states is actually lifted by quantum electrodynamic effects. The tiny upward shift of the $^2S_{1/2}$ state is called the Lamb shift.

It follows that the energy shift is

$$\Delta E = \frac{1}{4} m_e c^2 (Z\alpha)^4 \frac{\begin{Bmatrix} \ell \\ -\ell - 1 \end{Bmatrix}}{n^3 l(l + 1/2)(l + 1)}$$

valid for $l \neq 0$. When the effects of H_1 and H_2 are combined we get

$$\Delta E = \frac{1}{4} m_e c^2 (Z\alpha)^4 \left(\frac{1}{j + 1/2} - \frac{3}{4n} \right) \quad (12-16)$$

valid for both values of $l = j \pm 1/2$ and also for $l = 0$.

The splitting is depicted graphically in Fig. 12-1. A very interesting result is that the corrections add up in a manner that leaves the $^2P_{1/2}$ and the $^2S_{1/2}$ states degenerate. A more careful discussion, using the relativistic Dirac equation, does not alter this result. In 1947, a very delicate microwave absorption experiment carried out by Lamb and Retherford showed that there was, indeed, a tiny splitting of the two levels. The magnitude of the splitting, of order $m_e c^2 (Z\alpha)^4 \alpha \log \alpha$, could be explained by the additional interaction of the electron with its own electromagnetic field—that is, as a self-energy effect. These matters are outside of the scope of this book.

12-3 THE ANOMALOUS ZEEMAN EFFECT

Particles of charge $-e$ and mass M that move in a magnetic field $\mathbf{B}$ experience a potential given by

$$H_B = -\boldsymbol{\mu} \cdot \mathbf{B} = \frac{e}{2M} \mathbf{L} \cdot \mathbf{B} \quad (12-17)$$

If we choose $\mathbf{B}$ to define the z -axis, then the perturbation gives rise to the following energy shifts, known as the *Zeeman effect*:

$$\begin{aligned} \Delta E_{nlm} &= \langle nlm | H_B | nlm \rangle = \frac{eB}{2M} \langle nlm | L_z | nlm \rangle \\ &= \frac{eB\hbar}{2M} m \end{aligned} \quad (12-18)$$

where $m = l, l - 1, l - 2, \dots, -l$. The $(2l + 1)$ degenerate levels are split by the magnetic field. The presence of even numbers of lines was a mystery, cleared up by the discovery of spin. Because of the spin, there is an additional intrinsic magnetic moment present, and the new perturbation for an electron is now

$$H_B = \frac{e}{2m_e} (\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B} \quad (12-19)$$

Here we have used the fact that the gyromagnetic ratio for the electron is $g = 2$. Since spin-orbit coupling is generally larger than the perturbation due to the magnetic field, we take our unperturbed Hamiltonian to be

$$H_0 = \frac{\mathbf{p}^2}{2m_e} - \frac{Ze^2}{4\pi\epsilon_0 r} + \frac{1}{2m_e^2 c^2} \frac{Ze^2}{4\pi\epsilon_0} \frac{\mathbf{L} \cdot \mathbf{S}}{r^3} \quad (12-20)$$

The only effect of this is that we must calculate the matrix elements of H_B with the eigenstates of $\mathbf{J}^2$ and J_z . If we again choose the z -axis to be defined by $\mathbf{B}$, we need to calculate

$$\frac{eB}{2m_e} \langle \phi_{jm,l} | L_z + 2S_z | \phi_{jm,l} \rangle = \frac{eB}{2m_e} \langle \phi_{jm,l} | J_z + S_z | \phi_{jm,l} \rangle \quad (12-21)$$

The first term is easy to evaluate. It gives $(eB\hbar/2m_e)m_j$.

The second term can be evaluated in one of several ways. There is the brute-force evaluation, in which one calculates the matrix elements using the eigenfunctions given in (10A-13) and (10A-14). The second method is more intuitive. It can be established quite rigorously (which we don't do here) and it has a more general applicability—say, to spins other than $s = 1/2$. The basic idea is this: The “vector” $\mathbf{J}$ is a constant of the motion, while $\mathbf{S}$ and $\mathbf{L}$ are not. These vectors each have a fixed length, so that they precess about the direction of $\mathbf{J}$ (Fig. 12-2). The components of $\mathbf{S}$ and $\mathbf{L}$ along $\mathbf{J}$ are fixed, while the other components, rotating as they do about the $\mathbf{J}$ axis, average out to zero. This leads to the rule

$$\mathbf{S} \rightarrow \mathbf{J} \frac{(\mathbf{S} \cdot \mathbf{J})}{\mathbf{J}^2} = \mathbf{J} \frac{(\mathbf{S} \cdot \mathbf{J})}{\hbar^2 j(j+1)} \quad (12-22)$$

In the evaluation of $\langle \phi_{jm,l} | S_z | \phi_{jm,l} \rangle$ we therefore need to calculate $\mathbf{S} \cdot \mathbf{J}$, and this is easily done using $\mathbf{L} = \mathbf{J} - \mathbf{S}$. In fact,

$$\mathbf{S} \cdot \mathbf{J} = \frac{1}{2} [\mathbf{J}^2 + \mathbf{S}^2 - \mathbf{L}^2] = \frac{\hbar^2}{2} [j(j+1) + \frac{3}{4} - l(l+1)]$$

The above means that we replace $\langle S_z \rangle$ by

$$\langle J_z \rangle \frac{j(j+1) - l(l+1) + \frac{3}{4}}{2j(j+1)} \quad (12-23)$$

Adding the two terms yields

$$\Delta E_B = \frac{eB\hbar}{2m_e} m_j \left(1 + \frac{j(j+1) - l(l+1) + \frac{3}{4}}{2j(j+1)} \right) \quad (12-24)$$

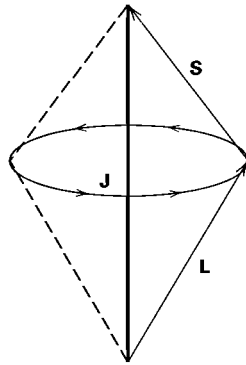


Figure 12-2 Sketch showing the precession of the orbital angular momentum vector and the spin vector about the constant total angular momentum vector.

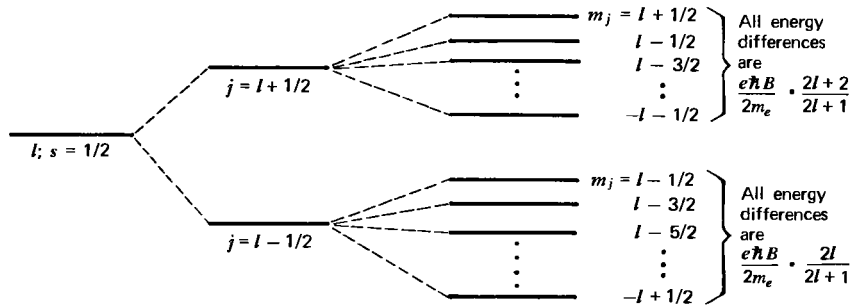


Figure 12-3 General representation of anomalous Zeeman effect.

This can be evaluated for $j = l + 1/2$ and for $j = l - 1/2$, and yields

$$\Delta E_B = m_j \left(\frac{eB\hbar}{2m_e} \right) \left(1 \pm \frac{1}{2l + 1} \right) \quad (12-25)$$

for $j = l \pm 1/2$. The splitting is depicted in Fig. 12-3. The lines drawn obey the so-called *selection rules*, $\Delta m_j = 0, \pm 1$.

The splitting between the levels is not the same for every multiplet, so that the pattern is complicated. For example for $n = 2$, the $^2P_{3/2}$ state splits into four lines, with the splitting twice as large as of the two states in the $^2P_{1/2}$ levels (Fig. 12-4). If the external field is strong enough so that the spin-orbit coupling can be neglected, we may use the ordinary

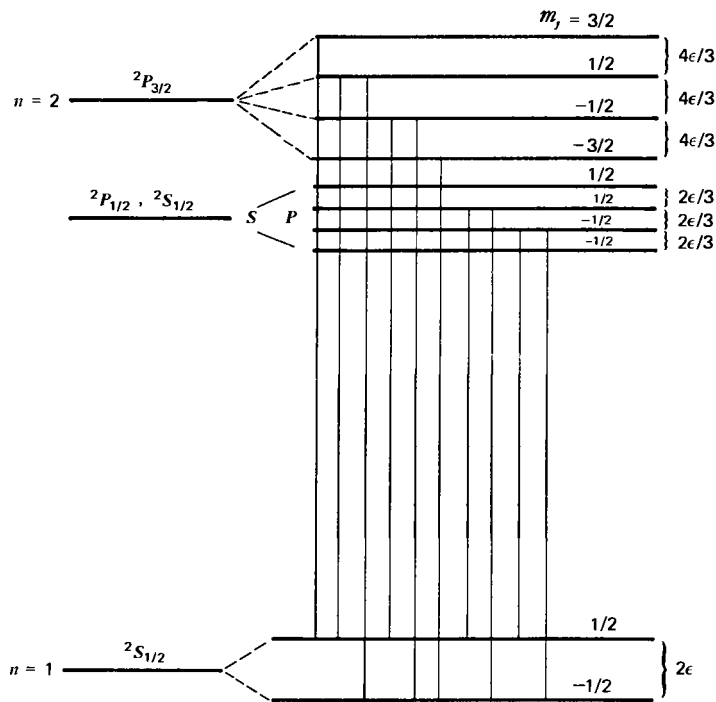


Figure 12-4 Zeeman effect in hydrogen, ϵ represents the energy $e\hbar B/2m_e$. The transitions for which $l = 1$, $\Delta m = 1, 0, -1$ are drawn in the figure. The location of the unperturbed states is given by Fig. 12-1.

hydrogenic wave functions simply multiplied by spinors—that is, eigenstates of $\mathbf{L}^2$, L_z , $\mathbf{S}^2$, and S_z . If we call the eigenvalues of L_z and S_z $\hbar m_l$, $\hbar m_s$, then

$$\Delta E_B = \frac{eB\hbar}{2m_e} (m_l + 2m_s) \quad (12-26)$$

Here the $n = 2$, $l = 1$ states are split into five levels, corresponding to $m_l = 1, 0, -1$ and $m_s = 1/2, -1/2$.

12-4 HYPERFINE STRUCTURE

In addition to the *fine structure* caused by the spin-orbit coupling, there is a very tiny *hyperfine splitting*, which is really a permanent Zeeman effect due to the magnetic field generated by the magnetic dipole moment of the nucleus. If the spin operator of the nucleus is denoted by $\mathbf{I}$, then the magnetic dipole operator is

$$\mathbf{M} = \frac{Ze g_N}{2M_N} \mathbf{I} \quad (12-27)$$

where Z is the nuclear charge, M_N the nuclear mass, and g_N its gyromagnetic ratio. Such a dipole gives rise to a magnetic field. We take the result from any textbook on electromagnetism. If the dipole is placed at the origin, the field at a point $\mathbf{r}$ is

$$\begin{aligned} \mathbf{B} &= \frac{\mu_0}{4\pi} \left(\frac{3\mathbf{r}(\mathbf{r} \cdot \mathbf{M}) - r^2 \mathbf{M}}{r^5} + \frac{8\pi}{3} \mathbf{M} \delta(\mathbf{r}) \right) \\ &= \frac{Ze g_N \mu_0}{2M_N 4\pi} \left(\frac{3\mathbf{r}(\mathbf{r} \cdot \mathbf{I}) - r^2 \mathbf{I}}{r^5} + \frac{8\pi}{3} \mathbf{I} \delta(\mathbf{r}) \right) \end{aligned} \quad (12-28)$$

Thus the hyperfine perturbation is

$$\begin{aligned} H_{hf} &= \frac{e}{2m_e} (\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B} \\ &= \frac{Ze^2 g_N \mu_0}{4M_N m_e 4\pi} \left(\frac{3\mathbf{r}(\mathbf{r} \cdot \mathbf{I}) - r^2 \mathbf{I}}{r^5} + \frac{8\pi}{3} \mathbf{I} \delta(\mathbf{r}) \right) \cdot (\mathbf{L} + 2\mathbf{S}) \end{aligned}$$

With $\mu_0 = \frac{1}{\epsilon_0 c^2}$, we get

$$H_{hf} = \frac{Ze^2}{4\pi\epsilon_0} \frac{g_N}{4M_N m_e c^2} \left(\frac{3\mathbf{r}(\mathbf{r} \cdot \mathbf{I})}{r^5} + \frac{8\pi}{3} \mathbf{I} \delta(\mathbf{r}) \right) \cdot (\mathbf{L} + 2\mathbf{S}) \quad (12-29)$$

An order of magnitude estimate yields

$$\begin{aligned} \frac{Ze^2}{4\pi\epsilon_0 \hbar c} \frac{\hbar c g_N}{4M_N m_e c^2} \frac{\hbar^2}{a_0^3} &\approx Z\alpha \frac{\hbar^3 c}{M_N m_e} \left(\frac{Z\alpha m_e}{\hbar} \right)^3 \\ &\approx (Z\alpha)^4 m_e c^2 \frac{m_e}{M_N} \end{aligned} \quad (12-30)$$

This is a factor m_e/M_N smaller than the typical spin-orbit coupling.

For a detailed calculation we shall only be interested in the hyperfine structure of an $l = 0$ state of a hydrogenlike atom. We may therefore ignore the operator $\mathbf{L}$ and we are left with

$$H_{hf} = \frac{Ze^2}{4\pi\epsilon_0} \frac{g_N}{2M_N m_e c^2} \left(\frac{3(\mathbf{r} \cdot \mathbf{I})(\mathbf{r} \cdot \mathbf{S})}{r^5} + \frac{8\pi}{3} \mathbf{I} \cdot \mathbf{S} \delta(\mathbf{r}) \right) \quad (12-31)$$

We must evaluate

$$\int d^3r |\psi_{n00}(\mathbf{r})|^2 \left(\frac{3(\mathbf{r} \cdot \mathbf{I})(\mathbf{r} \cdot \mathbf{S})}{r^5} + \frac{8\pi}{3} \mathbf{I} \cdot \mathbf{S} \delta(\mathbf{r}) \right) \quad (12-32)$$

A great simplification arises because $|\psi_{n00}(\mathbf{r})|^2$ is spherically symmetric. In doing the angular integration we have, for example,

$$\int d\Omega (\mathbf{r} \cdot \mathbf{S})(\mathbf{r} \cdot \mathbf{I}) = \int d\Omega (xS_x + yS_y + zS_z)(xI_x + yI_y + zI_z)$$

Now

$$\int d\Omega (xy) = \int d\Omega (xz) = \int d\Omega (yz) = 0$$

and

$$\int d\Omega x^2 = \int d\Omega y^2 = \int d\Omega z^2 = \frac{1}{3} \int d\Omega (x^2 + y^2 + z^2) = \frac{1}{3} \int d\Omega r^2 = \frac{4\pi}{3} r^2$$

We are thus left with $r^2 \mathbf{S} \cdot \mathbf{I}$, and we see that the first two terms cancel. We are left with the last term. This gives

$$\int d^3r |\psi_{n00}(\mathbf{r})|^2 \delta(\mathbf{r}) = \frac{1}{4\pi} |R_{n0}(0)|^2 = \frac{1}{\pi} \left(\frac{Z\alpha m_e c}{\hbar} \right)^3 \frac{1}{n^3} \quad (12-33)$$

For hydrogen in the ground state $Z = 1$, $n = 1$ and we get

$$\frac{e^2}{4\pi\epsilon_0} \frac{g_P}{2M_P m_e c^2} \frac{8\pi}{3} \mathbf{I} \cdot \mathbf{S} \frac{1}{\pi} \left(\frac{m_e \alpha c}{\hbar} \right)^3 = \frac{4}{3} \alpha^4 m_e c^2 g_P \left(\frac{m_e}{M_P} \right) \frac{\mathbf{I} \cdot \mathbf{S}}{\hbar^2} \quad (12-34)$$

As a last step we introduce the total spin $\mathbf{F} = \mathbf{S} + \mathbf{I}$. Then

$$\frac{\mathbf{S} \cdot \mathbf{I}}{\hbar^2} = \frac{1}{2} [F(F+1) - S(S+1) - I(I+1)] \quad (12-35)$$

For $F = 1$ this is $1/4$; for $F = 0$, it equals $-3/4$. This leads to the final answer

$$\Delta E_{hf} = \frac{4}{3} \alpha^4 m_e c^2 \left(\frac{m_e}{M_P} \right) g_P \quad (12-36)$$

The proton gyromagnetic ratio is $g_P = 5.56$, so that the frequency of the radiation emitted in the transition between the two levels is $\nu \approx 1420$ MHz. This corresponds to the wavelength of $\lambda = 21.1$ cm.

The radiation arising from this transition plays an important role in astronomy. In a gas of neutral atoms, the $F = 1$ state cannot be excited by ordinary radiation, because of a selection rule that strongly suppresses transitions in which there is no change in orbital angular momentum. Both the $F = 1$ and the $F = 0$ states have zero angular momentum. On the other hand, there are other mechanisms that can cause transitions. The $F = 1$ state can, for example, be excited by collisions, and the return to the $F = 0$ ground state can be detected. From an analysis of the intensity of the 21-cm radiation received, astronomers have learned a great deal about the density distribution of neutral hydrogen in interstellar space, as well as the motion and the temperature of the gas clouds containing the hydrogen. The average number of neutral hydrogen atoms appears to be about 1 cm^{-3} in the galactic plane near the sun, and the temperature is of the order of 100 K.

12-5 COMMENTS ON REDUCED MASS EFFECTS

We ignored reduced mass effects in the perturbation calculations because of the small magnitude of m_e/M_p . In the past thirty years it has been possible to study spectra of unstable atoms such as atoms in which a short-lived μ^- (a heavy electron of mass $m_\mu = 205.8 m_e$) forms a bound state with a proton, or positronium in which the e^- forms a bound state with the positron e^+ , a particle identical to the electron, except for the sign of its charge (and magnetic moment). In these cases the reduced mass effects are more important. We briefly discuss how these affect the discussion of the relativistic effects. We shall do this in the language of the hydrogen atom.

If we take into account the relativistic kinetic energy of the proton as well as the electron, then the corrections to the nonrelativistic kinetic energy can be written as

$$H_1 = -\frac{1}{8} \frac{(\mathbf{p}^2)^2}{c^2} \left(\frac{1}{m_e^3} + \frac{1}{M_p^3} \right)$$

which, after a little algebra, can be written as

$$H_1 = \frac{1}{8} \frac{(\mathbf{p}^2)^2}{\mu m_e^2 c^2} \left(1 - \frac{m_e}{M_p} + \left(\frac{m_e}{M_p} \right)^2 \right) \quad (12-37)$$

In the spin-orbit term $\mathbf{M}$ is unaltered, but $\mathbf{v} = \mathbf{p}/\mu$.

In our discussion of the anomalous Zeeman effect, the perturbation now reads

$$H_3 = \left(\frac{e}{2\mu} \mathbf{L} + \frac{eg}{2m_e} \mathbf{S} \right) \cdot \mathbf{B} \quad (12-38)$$

and (12-21) becomes

$$\begin{aligned} \left\langle \phi_{jm,l} \left| \frac{eB}{2\mu} L_z + \frac{egB}{2m_e} S_z \right| \phi_{jm,l} \right\rangle &= B \left\langle \phi_{jm,l} \left| \frac{e}{2\mu} J_z + \left(\frac{eg}{2m_e} - \frac{e}{2\mu} \right) S_z \right| \phi_{jm,l} \right\rangle \\ &= \frac{eB}{2\mu} \left[\hbar m_j + \left(g \frac{\mu}{m_e} - 1 \right) \langle \phi_{jm,l} | S_z | \phi_{jm,l} \rangle \right] \end{aligned} \quad (12-39)$$

Reduced mass effects can be ignored in our discussion of hyperfine structure for hydrogen, since the whole effect is a factor of m_e/M_N smaller than the atomic effects under consideration.

Numerically, the reduced mass effects are small. The table that follows² lists the energy shifts in millielectronvolts including relativistic and spin-orbit terms:

Level	Energy Shift with $\mu = m_e$	Energy Shift Including Red. Mass Effects
1 $S_{1/2}$	-0.18113	-0.18074
2 $S_{1/2}$	-0.05660	-0.05648
2 $P_{1/2}$	-0.05660	-0.05651
2 $P_{3/2}$	-0.01132	-0.01128

The reduced mass effects on the energy shifts are of the order of 4×10^{-7} eV. Quantum electrodynamic effects (Lamb shift) are of the order of 4×10^{-6} eV. Thus for hydrogen there is no point in keeping the reduced mass effects, unless one is interested in very high precision calculations. On the other hand, for a system like positronium, the $e^- - e^+$ bound state, reduced mass effects are crucial, since the two particles have the same mass. There the hyperfine splitting is not suppressed relative to the spin-orbit or relativistic effect.

PROBLEMS

1. If the general form of the spin-orbit coupling for a particle of mass m and spin $\mathbf{S}$ moving in a potential $V(r)$ is

$$H_{so} = \frac{1}{2m^2c^2} \mathbf{S} \cdot \mathbf{L} \frac{1}{r} \frac{dV(r)}{dr}$$

what is the effect of that coupling on the spectrum of a three-dimensional harmonic oscillator?

(Note: The unperturbed SHO spectrum is $\hbar\omega(2n_r + l + 3/2)$, where $n_r = 0, 1, 2, \dots$, and l is the orbital angular momentum.)

2. Calculate the energy spectrum of the $n = 2$ states in the real hydrogen atom, ignoring hyperfine structure. How is the spectrum changed when the atom is placed in a magnetic field of 2.5T?
3. Consider hydrogen gas in its ground state. What is the effect of a magnetic field on the hyperfine structure? Calculate the spectrum for $B = 10^{-4}$ T and for $B = 1$ T. [Hint: To solve this, set up an eigenvalue problem for the interaction

$$A(\mathbf{S} \cdot \mathbf{I})/\hbar^2 + a(S_z/\hbar) + b(I_z/\hbar)$$

with

$$A = g_F(4m_e/3M_p)\alpha^4 m_e c^2; a = e\hbar B/m_e; b = -eg_F\hbar B/2M_p$$

4. Consider the transitions $3^2P_{1/2} \rightarrow 1^2S_{1/2}$; $3^2P_{3/2} \rightarrow 1^2S_{1/2}$ in hydrogenlike atoms.
 - (a) Calculate the wavelengths of light emitted in these transitions as a function of Z .
 - (b) In sodium the wavelengths are 589.592 nm and 588.995 nm, respectively. What value of Z does this correspond to?
5. Consider a harmonic oscillator in three dimensions. If the relativistic expression for the kinetic energy is used, what is the shift in the ground-state energy?

²The table was kindly provided by Professor J. S. Tenn.

6. The deuteron consists of a proton (charge $+e$) and a neutron (charge 0) in a state of total spin 1 and total angular momentum $J = 1$. The g -factors for the proton and neutron are

$$g_P = 2(2.7896)$$

$$g_N = 2(-1.9103)$$

- (a) What are the possible orbital angular momentum states for this system? If it is known that the state is primarily ${}^3S_1$, what admixture is allowed given that parity is conserved?
- (b) Write an expression for the interaction of the deuteron with an external magnetic field and calculate the Zeeman splitting. Show that if the interaction with the magnetic field is written in the form

$$V = -\boldsymbol{\mu}_{\text{eff}} \cdot \mathbf{B}$$

then the effective magnetic moment of the deuteron is the sum of the proton and neutron magnetic moments, and any deviation from that result is due to an admixture of non- S state to the wave function.

Chapter 13

Many-Particle Systems

Our discussion of the quantum mechanics of a single particle is easily generalized to an N -particle system. For such a system in one space dimension, the particles are described by a wave function $\psi(x_1, x_2, x_3, \dots, x_N)$, normalized such that

$$\int \int \dots \int dx_1 dx_2 \dots dx_N |\psi(x_1, x_2, x_3, \dots, x_N)|^2 = 1 \quad (13-1)$$

The interpretation of $|\psi(x_1, x_2, x_3, \dots, x_N)|^2$ is a generalization of the interpretation of the one-particle wave function: It yields the probability density for finding particle 1 near x_1 , particle 2 near x_2 , $\dots$, particle N near x_N . The time evolution of such a wave function is given by the solution of the partial differential equation

$$i\hbar \frac{\partial \psi(x_1, x_2, x_3, \dots, x_N)}{\partial t} = H\psi(x_1, x_2, x_3, \dots, x_N) \quad (13-2)$$

where the Hamiltonian H is again constructed in correspondence with the classical form of the energy

$$H = \sum_{i=1}^N \frac{p_i^2}{2m_i} + V(x_1, x_2, x_3, \dots, x_N) \quad (13-3)$$

as

$$H = -\hbar^2 \left(\frac{1}{2m_1} \frac{\partial^2}{\partial x_1^2} + \frac{1}{2m_2} \frac{\partial^2}{\partial x_2^2} + \dots + \frac{1}{2m_N} \frac{\partial^2}{\partial x_N^2} \right) + V(x_1, x_2, \dots, x_N) \quad (13-4)$$

The whole formalism of quantum mechanics developed up to here is easily generalized, with the proviso that operators describing single-particle observables commute when they refer to different particles; for example,

$$[p_i, x_j] = -i\hbar \delta_{ij} \quad (13-5)$$

13-1 THE TWO-PARTICLE SYSTEM

It is convenient to discuss this in three-dimensional space. This extension is trivial. We begin by considering a two-particle system described by the Hamiltonian

$$H = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + V(\mathbf{r}_1, \mathbf{r}_2) \quad (13-6)$$

Let us change variables to

$$\begin{aligned} \mathbf{P} &= \mathbf{p}_1 + \mathbf{p}_2 \\ \mathbf{r} &= \mathbf{r}_1 - \mathbf{r}_2 \end{aligned} \quad (13-7)$$

and introduce the center of mass coordinate

$$\mathbf{R} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{M} \quad (13-8)$$

where $M = m_1 + m_2$ is the total mass. We see that

$$[P_a, R_b] = -i\hbar \delta_{ab} \quad (a, b = 1, 2, 3) \quad (13-9)$$

The variable conjugate to the relative distance $\mathbf{r}$ is

$$\mathbf{p} = \frac{m_2 \mathbf{p}_1 - m_1 \mathbf{p}_2}{M} \quad (13-10)$$

We solve to get $\mathbf{p}_1 = \mathbf{p} + (m_1/M)\mathbf{P}$ and $\mathbf{p}_2 = -\mathbf{p} + (m_2/M)\mathbf{P}$ from which we can calculate the kinetic energy in the form

$$\frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} = \frac{\mathbf{P}^2}{2M} + \frac{1}{2}\mathbf{p}^2 \left(\frac{1}{m_1} + \frac{1}{m_2} \right) = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2\mu} \quad (13-11)$$

where μ is the reduced mass. In writing the eigenvalue equation

$$\left(\frac{\mathbf{P}_{op}^2}{2\mu} + \frac{\mathbf{P}_{op}^2}{2M} + V(\mathbf{r}, \mathbf{R}) \right) \psi(\mathbf{r}, \mathbf{R}) = E \psi(\mathbf{r}, \mathbf{R}) \quad (13-12)$$

we treat $\mathbf{P}_{op}$ and $\mathbf{p}_{op}$ as operators. If the potential only depends on the separation between the variables, so that $V = V(\mathbf{r})$, we may write the wave function as a product of the eigenfunction of $\mathbf{P}_{op}$ and $u(\mathbf{r})$, where

$$\psi(\mathbf{r}, \mathbf{R}) = e^{i\mathbf{P}\cdot\mathbf{R}/\hbar} u(\mathbf{r}) \quad (13-13)$$

Here $\mathbf{P}$ is the eigenvalue of $\mathbf{P}_{op}$ and $u(\mathbf{r})$ is a solution of

$$\left(\frac{\mathbf{P}_{op}^2}{2\mu} + V(\mathbf{r}) \right) u(\mathbf{r}) = \left(E - \frac{\mathbf{P}^2}{2M} \right) u(\mathbf{r}) \quad (13-14)$$

This is a one-particle equation for a particle with reduced mass μ , moving in the potential $V(\mathbf{r})$, and energy E reduced by the kinetic energy of the two-particle system. We have already used this result in dealing with the hydrogen atom as a single-particle system with reduced mass for the electron.

We may ask, under what circumstances does V depend on $\mathbf{r}$ only? This happens when the Hamiltonian has a special *symmetry* property, which is that *it is unchanged by the uniform displacement of all particles*. If H is unchanged when

$$\mathbf{r}_i \rightarrow \mathbf{r}_i + \mathbf{a} \quad (13-15)$$

where $\mathbf{a}$ is the same for all particles, then V can only depend on the separation of two particles. The kinetic energy involves derivatives with respect to the $\mathbf{r}_i$ so that it also obeys the above rule.

We show in Supplement 13-A [www.wiley.com/college/gasiorowicz] that the invariance of the Hamiltonian under the transformation (13-15) leads to the existence of a conserved quantity, which turns out to be the *total momentum*. This is very similar to the result that the conservation of total angular momentum is a consequence of the invariance of the Hamiltonian under rotations.

13-2 IDENTICAL PARTICLES

There is compelling evidence that electrons are indistinguishable from one another. If this were not so, then the spectrum of an atom—helium, say—would vary from experiment to experiment, depending on “what kind” of electrons were contained in the sample under investigation. In fact, no such variation has ever been observed. Similarly molecular and nuclear spectra are always the same—indication that protons are indistinguishable, as are neutrons. Similar evidence from high-energy physics indicates very strongly that other particles—for example, *pions*—are indistinguishable. The same holds for photons, an important property utilized in lasers.

Indistinguishability is a purely quantum mechanical property: In classical mechanics it is in principle possible to follow the orbits of all particles, so that they are never really indistinguishable. We have learned that electrons are characterized by an internal quantum number, its *spin*. Thus a complete set of quantum numbers for the description of an electron must include the spin label. This is generally chosen to be the value of the z -component of the spin, which has the value $\pm\hbar/2$. This spin state will be labeled by¹ σ , which will be double-valued. What this means is that two electrons, identical in every respect (except for the spin), can still be distinguished by their σ -value. A third electron with the same quantum numbers as the other two must have a spin label that is identical to that of at least one of the other electrons. The existence of the spin label has a further effect on the consequences of indistinguishability, which we discuss next.

The Exchange Operator

A Hamiltonian for indistinguishable particles must be completely symmetric in the coordinates of the particles. For a two-particle system, if the potential does not depend on the spin labels, the Hamiltonian is

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + V(x_1, x_2) \quad (13-16)$$

$$V(x_1, x_2) = V(x_2, x_1) \quad (13-17)$$

We write this symmetry symbolically as

$$H(1, 2) = H(2, 1) \quad (13-18)$$

and it is understood that if the Hamiltonian contains operators that refer to the spins of the two particles, then these too must be included in the labeling “1,” “2.”

A wave function for an N -particle system, labeled by the total energy, for example, now needs to be labeled by the spin labels $\sigma_1, \sigma_2, \dots, \sigma_N$. For a two-particle system, the energy eigenvalue equation now reads

$$H(1, 2)u_{E\sigma_1\sigma_2}(1, 2) = Eu_{E\sigma_1\sigma_2}(1, 2) \quad (13-19)$$

¹The spin index is sometimes labeled by $\pm$, or even by the “up” and “down” symbols $\uparrow, \downarrow$.

Since the labeling does not matter, we may, by interchanging the “1” and “2” labels, write this as

$$H(2, 1)u_{E\sigma_2\sigma_1}(2, 1) = Eu_{E\sigma_2\sigma_1}(2, 1) \quad (13-20)$$

On the other hand, using (13-18) we also have

$$H(1, 2)u_{E\sigma_2\sigma_1}(2, 1) = Eu_{E\sigma_2\sigma_1}(2, 1) \quad (13-21)$$

Let us now use the formal approach that we used in our discussion of parity. We introduce an *exchange operator* P_{12} , which, acting on a state, interchanges the two identical particles that we labeled by 1 and 2. Thus

$$P_{12}u_{E\sigma_1\sigma_2}(1, 2) = u_{E\sigma_2\sigma_1}(2, 1) \quad (13-22)$$

Equation (13-21) may be written as follows:

$$\begin{aligned} H(1, 2)P_{12}u_{E\sigma_1\sigma_2}(1, 2) &= Eu_{E\sigma_2\sigma_1}(2, 1) \\ &= EP_{12}u_{E\sigma_1\sigma_2}(1, 2) \\ &= P_{12}Eu_{E\sigma_1\sigma_2}(1, 2) \\ &= P_{12}H(1, 2)u_{E\sigma_1\sigma_2}(1, 2) \end{aligned} \quad (13-23)$$

Since this applies for the complete set of $u_{E\sigma_1\sigma_2}(1, 2)$ of simultaneous eigenfunctions of H and the spin operators, we deduce the *operator relation*

$$[H, P_{12}] = 0 \quad (13-24)$$

Thus P_{12} , likely parity, is a constant of the motion. Further, since two exchanges $1 \rightarrow 2$, $2 \rightarrow 1$ bring back the original state, we have

$$(P_{12})^2 = 1 \quad (13-25)$$

so that the eigenvalues of P_{12} are ± 1 . Just as even and odd functions are eigenfunctions of the parity operator, here the eigenstates are the symmetric and antisymmetric combinations

$$\psi^{(S)}(1, 2) = \frac{1}{N_{2S}} [\psi(1, 2) + \psi(2, 1)] \quad (13-26)$$

$$\psi^{(A)}(1, 2) = \frac{1}{N_{2A}} [\psi(1, 2) - \psi(2, 1)] \quad (13-27)$$

where the N_2 are normalization constants. The fact that P_{12} is a constant of the motion implies that a state that is symmetric at an initial time will always be symmetric, and an antisymmetric state will always be antisymmetric.

13-3 THE PAULI PRINCIPLE

It is an important *law of nature* that the symmetry or antisymmetry under the interchange of two particles is a characteristic of the particles, and not something that can be arranged in the preparation of the initial state. The law, which was discovered by Pauli, states that

1. Systems consisting of identical particles of half-odd-integral spin (i.e., spin $1/2$, $3/2$, ...) are described by antisymmetric wave functions. Such particles are called *fermions*, and are said to obey Fermi-Dirac statistics.
2. Systems consisting of identical particles of integral spin (spin 0 , 1 , 2 , ...) are described by symmetric wave functions. Such particles are called *bosons*, and are said to obey Bose-Einstein statistics. We shall primarily be concerned with elec-

trons, protons, and neutrons, which have spin 1/2, and with spin 0 bosons, which carry no spin labeling.

The law extends to N -particle states. For a system of N identical fermions, the wave function is antisymmetric under the interchange of *any* pair of particles. For example, a three-particle wave function, properly antisymmetrized, has the form

$$\begin{aligned} \psi^{(A)}(1, 2, 3) = \frac{1}{N_{3A}} [\psi(1, 2, 3) - \psi(2, 1, 3) + \psi(2, 3, 1) \\ - \psi(3, 2, 1) + \psi(3, 1, 2) - \psi(1, 3, 2)] \end{aligned} \quad (13-28)$$

whereas the three identical boson wave function has the form

$$\begin{aligned} \psi^{(S)}(1, 2, 3) = \frac{1}{N_{3S}} [\psi(1, 2, 3) + \psi(2, 1, 3) + \psi(2, 3, 1) \\ + \psi(3, 2, 1) + \psi(3, 1, 2) + \psi(1, 3, 2)] \end{aligned} \quad (13-29)$$

It should be stressed that for more than two identical particles, one could in principle have mixed symmetries; for example, the wave function is antisymmetric under the exchange of (1, 2) and (1, 3), but symmetric under the exchange (2, 3). The Pauli principle excludes such mixed symmetry states.

N Fermions in a Potential Well

Let us now consider a very interesting special case, in which N fermions do not interact with each other, but do interact with a common potential. In that case

$$H = \sum_{i=1}^N H_i \quad (13-30)$$

where

$$H_i = \frac{p_i^2}{2m} + V(x_i) \quad (13-31)$$

The eigenstates of the one-particle Hamiltonian are denoted by $u_{E\sigma_k}(x_k)$, where

$$H_k u_{E\sigma_k}(x_k) = E_k u_{E\sigma_k}(x_k) \quad (13-32)$$

It is understood that for each value of E_k there are two possible values of the spin σ_k label.

A solution of

$$H u_E(1, 2, \dots, N) = E u_E(1, 2, \dots, N) \quad (13-33)$$

is

$$u_E(1, 2, 3, \dots, N) = u_{E_1\sigma_1}(x_1) u_{E_2\sigma_2}(x_2), \dots, u_{E_N\sigma_N}(x_N) \quad (13-34)$$

We shall write this as

$$u_E(1, 2, 3, \dots, N) = u_{E_1}(x_1) u_{E_2}(x_2), \dots, u_{E_N}(x_N) \quad (13-35)$$

where we have suppressed the σ_i labels that go with the E_i . Also

$$E = E_1 + E_2 + \dots + E_N \quad (13-36)$$

Our task now is to antisymmetrize (13-35). If there are only two particles, we evidently have

$$u^{(A)}(1, 2) = \frac{1}{\sqrt{2}} [u_{E_1}(x_1) u_{E_2}(x_2) - u_{E_1}(x_2) u_{E_2}(x_1)] \quad (13-37)$$

With three particles, the form is

$$u^{(A)}(1, 2, 3) = \frac{1}{\sqrt{6}} [u_{E_1}(x_1)u_{E_2}(x_2)u_{E_3}(x_3) - u_{E_1}(x_2)u_{E_2}(x_1)u_{E_3}(x_3) \\ + u_{E_1}(x_2)u_{E_2}(x_3)u_{E_3}(x_1) - u_{E_1}(x_3)u_{E_2}(x_2)u_{E_3}(x_1) \\ + u_{E_1}(x_3)u_{E_2}(x_1)u_{E_3}(x_2) - u_{E_1}(x_1)u_{E_2}(x_3)u_{E_3}(x_2)] \quad (13-38)$$

For N particles, the answer is a determinant, the so-called *Slater determinant*:²

$$u^{(A)}(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} u_{E_1}(x_1) & u_{E_1}(x_2) & \cdots & u_{E_1}(x_N) \\ u_{E_2}(x_1) & u_{E_2}(x_2) & \cdots & u_{E_2}(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ u_{E_N}(x_1) & u_{E_N}(x_2) & \cdots & u_{E_N}(x_N) \end{vmatrix} \quad (13-39)$$

In the preceding three equations we have suppressed the label σ_k that goes with each E_k . Clearly the interchange of two particles involves the interchange of two columns in the determinant, and this changes the sign. If two electrons are in the same energy eigenstate—for example, $E_1 = E_2$ —and if they are in the same spin state—that is, the spin labels are the same $\sigma_1 = \sigma_2$ —then the determinant vanishes when $x_1 = x_2$; that is, the electrons cannot be at the same place. Thus the requirement of antisymmetry introduces an effective interaction between two fermions; qualitatively we see that two particles in the same state tend to stay away from each other, since the joint wave function vanishes when their separation goes to zero. Thus even noninteracting particles behave as if there were a repulsive interaction between them.

When Is Antisymmetrization Necessary?

The statement “no two electrons can be in the same quantum state” implies that a wave function for a system containing two electrons must be antisymmetric in the coordinates of the two electrons. The question arises whether we really have to worry about this when we consider a hydrogen atom on earth and another one on the moon. If they are both in the ground state, do the electrons necessarily have to have opposite spin states? What then happens when we consider a third hydrogen atom in its ground state? Intuition tells us that we do not have to worry, and intuition is correct. We can see this by examining whether there is a difference between using the totally uncorrelated wave function for two electrons

$$\psi_a(x_1)\psi_b(x_2) \quad (13-40)$$

and the antisymmetrized wave function

$$\frac{1}{\sqrt{2}} (\psi_a(x_1)\psi_b(x_2) - \psi_a(x_2)\psi_b(x_1)) \quad (13-41)$$

The normalization factor N is given by the requirement that

$$\frac{1}{N^2} \int dx_1 \int dx_2 |\psi_a(x_1)\psi_b(x_2) - \psi_a(x_2)\psi_b(x_1)|^2 = 1 \quad (13-42)$$

which with

$$\int dx |\psi_a(x)|^2 = \int dx |\psi_b(x)|^2 = 1 \quad (13-43)$$

²The wave function for N identical bosons is totally symmetric, and the general form is obtained by expanding the determinant in (13-39) and making all the signs positive.

leads to

$$\begin{aligned} N^2 &= 2 \left(1 - \left| \int dx \psi_a^*(x) \psi_b(x) \right|^2 \right) \\ &\equiv 2(1 - |S_{ab}|^2) \end{aligned} \quad (13-44)$$

Suppose we want to calculate the probability that the electron with the a label is in some spatial region R . For the uncorrelated wave function, which is $\Psi(x, y) = \psi_a(x)\psi_b(y)$, the probability density is given by

$$P(R) = \int_R dx \int dy |\psi_a(x)|^2 |\psi_b(y)|^2 = \int_R dx |\psi_a(x)|^2 \quad (13-45)$$

We integrate over the whole range of the coordinates of the b electron, since we do not care where it is.

For the antisymmetrized wave function we have

$$|\Psi(x, y)|^2 = \frac{1}{N^2} [\psi_a^*(x)\psi_b^*(y) - \psi_b^*(x)\psi_a^*(y)][\psi_a(x)\psi_b(y) - \psi_b(x)\psi_a(y)]$$

This is to be integrated over the domain R for those variables associated with the label a and over the whole range of coordinates associated with the b label. We thus get

$$\begin{aligned} P_a(R) &= \frac{1}{N^2} \int_R dx |\psi_a(x)|^2 \int_{-\infty}^{\infty} dy |\psi_b(y)|^2 \\ &\quad + \frac{1}{N^2} \int_R dy |\psi_a(y)|^2 \int_{-\infty}^{\infty} dx |\psi_b(x)|^2 \\ &\quad - \frac{1}{N^2} \int_R dx \int_R dy [\psi_a^*(x)\psi_b(x)\psi_b^*(y)\psi_a(y) + \psi_b^*(x)\psi_a(x)\psi_a^*(y)\psi_b(y)] \\ &= \frac{2}{N^2} \int_R dx |\psi_a(x)|^2 - \frac{2}{N^2} \int_R dx \int_R dy \psi_a^*(x)\psi_b(x)\psi_b^*(y)\psi_a(y) \end{aligned} \quad (13-46)$$

The interference term has both integrals over the range R , since the wave function for electron a appears in the x -region as well as the y -region. The difference will only be significant if the *overlap integral* $\int_R dx \psi_a^*(x)\psi_b(x)$ is important in the region R for the variable x . Since wave functions fall off exponentially for bound states, it is clear that this can only be important if the atoms are very near to each other.

For example, consider a two-electron system, each in a gaussian wave packet, one centered about the origin and the other about $x = L$. A calculation of the probability of finding an electron in a region R involves the overlap integral of $Ce^{-\beta x^2}$ and $Ce^{-\beta(x-L)^2}$. The overlap integral is of the form

$$C^2 \int_R dx e^{-\beta(x^2 + (x-L)^2)}$$

and this is easily seen to be proportional to $e^{-\beta L^2/2}$. Thus for L large, the overlap integral vanishes very rapidly, and our intuition that the wave function of an electron under consideration need not be antisymmetrized with any or all other, distant electrons, proves to be correct.

The Pauli exclusion principle must be taken into account in atoms and in molecules, but not in situations where the atoms are separated by significant distances. Even in crystal lattices, where the spacing between the atoms is several angstroms, the overlap is frequently small, and antisymmetrization is not necessary.

13-4 THE EXCLUSION PRINCIPLE AND THE TWO-PARTICLE PROBLEM

For identical particles $m_1 = m_2$, so that (13-10) shows that $\mathbf{p} = \mathbf{p}_1 - \mathbf{p}_2$. Thus, under the interchange of the two identical particles that obey (13-15), the kinetic energy does not change sign, while

$$V(\mathbf{r}) \rightarrow V(-\mathbf{r}) \quad (13-47)$$

The identity of the particles implies that when they are interchanged, the wave function can acquire a phase (for example change sign) but it must obey the same equation. This is only possible if $V(\mathbf{r}) = V(r)$; that is, the potential only depends on the *magnitude of the separation of the two particles*. Equation (13-14) is then one involving a central potential, and we can write the solution in the form

$$u(\mathbf{r}) = R_{nl}(r)Y_{lm}(\theta, \varphi) \quad (13-48)$$

An interchange of the two identical particles corresponds to the change

$$\begin{aligned} r &\rightarrow r \\ \theta &\rightarrow \pi - \theta \\ \varphi &\rightarrow \varphi + \pi \end{aligned} \quad (13-49)$$

The radial function remains unchanged, but under this transformation

$$Y_{lm}(\theta, \varphi) \rightarrow Y_{lm}(\pi - \theta, \varphi + \pi) = (-1)^l Y_{lm}(\theta, \varphi) \quad (13-50)$$

The two-electron system can be in a spin singlet state

$$\frac{1}{\sqrt{2}} (\chi_+^{(1)}\chi_-^{(2)} - \chi_-^{(1)}\chi_+^{(2)}) \quad (13-51)$$

which is antisymmetric under the interchange of the two electrons, or in the spin triplet state

$$\begin{aligned} &\chi_+^{(1)}\chi_+^{(2)} \\ &\frac{1}{\sqrt{2}} (\chi_+^{(1)}\chi_-^{(2)} + \chi_-^{(1)}\chi_+^{(2)}) \\ &\chi_-^{(1)}\chi_-^{(2)} \end{aligned} \quad (13-52)$$

which is symmetric under the interchange of the two electrons. Thus the exclusion principle, which demands that the total wave function be antisymmetric under the interchange of the two electrons, leads to the requirements that *spin singlet systems have $l = 0, 2, 4, 6, \dots$* , and *the spin triplet systems have $l = 1, 3, 5, \dots$* . We shall see an application of this when we discuss the helium atom.

A Digression on Parity

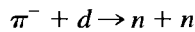
Note that the same argument can be applied to checking the properties of the Y_{lm} under inversion. The transformation $x \rightarrow -x$, $y \rightarrow -y$, and $z \rightarrow -z$ is identical to (13-49). Thus we see that a particle in an orbital angular momentum l state will have its wave function

changed by $(-1)^l$. Even l orbital states are also even-parity states, and odd l orbital states are odd-parity states. It should be noted, however, that the particles themselves can have an intrinsic parity. We can define the intrinsic parity of the electron and the proton and neutron to be even. In that case, the parity of a hydrogen $l = 1$ state, for example, is odd, while the parity of the ground state is even.

In relativistic quantum mechanics one can show that the intrinsic parity of an antiparticle of a fermion is opposite to that of the fermion. Thus the e^+ has negative intrinsic parity, and hence the ground state of positronium, for which $l = 0$, has *negative* parity.

An interesting application of these remarks occurs in elementary particle physics. One of the first unstable elementary particles to be discovered was the π meson predicted by Yukawa. This particle, which plays an important role in nuclear forces, comes in three charge states π^+ , π^0 , π^- . It was found to have spin 0, and the question arose whether the wave function of a pion—as this meson came to be called—was even or odd under reflection, assuming that the known particles, the proton and the neutron, had positive intrinsic parity. The following experiment was suggested.

Consider the capture of a π^- by a deuteron. A slow pion in liquid deuterium loses energy by a variety of mechanisms, until it finally ends up in the lowest Bohr orbit about the (pn) nucleus, and is then captured through the action of the nuclear forces. In the nuclear reaction



the angular momentum is 1; the pion has zero spin, the orbital angular momentum is zero in the lowest Bohr state, so that the only contribution is the angular momentum of the deuteron, which is 1. The two neutrons must therefore be in an angular momentum 1 state. If the total spin of the two neutrons is 0, then the orbital angular momentum must be 1. If the total spin of the two-neutron state is 1, then orbital angular momentum 0, 1, and 2 is possible, since adding two angular momenta of one unit each can yield 0, 1, and 2, and adding one unit to two units of angular momentum can yield 3, 2, and 1. However, a singlet state of two identical fermions must have even angular momentum, and is thus excluded. A triplet state must have odd orbital angular momentum, and this is possible if the orbital angular momentum is 1. Such a state, however, has odd parity by (13-50), and hence the pion must have odd parity.

13-5 THE EXCLUSION PRINCIPLE AND NONINTERACTING PARTICLES

The *exclusion principle* has huge consequences for the structure of atoms, as will be discussed in Chapter 14. We can get a sense of this effect by considering a very simple model of a many-electron system by considering N spin 1/2 particles that do not interact with each other, but each is placed in an infinite well. The potential is assumed to be a cubical box, with side L . In one dimension this problem was solved in Chapter 3. The eigenfunctions were found to be

$$u_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}$$

with energies $E_n = \frac{\hbar^2 \pi^2}{2mL^2} n^2$. The extension to three dimensions is straightforward, since the motion in the x , y , z directions are uncorrelated. Thus

$$u_E(x, y, z) = \left(\frac{2}{L}\right)^{3/2} \sin \frac{n_1 \pi x}{L} \sin \frac{n_2 \pi y}{L} \sin \frac{n_3 \pi z}{L} \quad (13-53)$$

with

$$E = \frac{\hbar^2 \pi^2}{2mL^2} (n_1^2 + n_2^2 + n_3^2) \quad (13-54)$$

Note that there is quite a lot of degeneracy in the problem: There are as many solutions for a given E as there are sets of integers $\{n_1, n_2, n_3\}$ that satisfy (13-54). Degeneracy is usually associated with the existence of mutually commuting observables and this example is no exception. In general when $H = H_x + H_y + H_z$ where

$$H_x = \frac{p_x^2}{2m} + V_1(x); H_y = \frac{p_y^2}{2m} + V_2(y); H_z = \frac{p_z^2}{2m} + V_3(z) \quad (13-55)$$

the eigenfunctions of H have the form

$$\psi_{E_1 E_2 E_3}(x, y, z) = u_{E_1}(x) v_{E_2}(y) w_{E_3}(z) \quad (13-56)$$

Where $u_{E_1}(x)$, $v_{E_2}(y)$, $w_{E_3}(z)$ are eigenfunction of H_x , H_y , and H_z respectively, with eigenvalues E_1 , E_2 and E_3 , and

$$E = E_1 + E_2 + E_3 \quad (13-57)$$

We may now ask for the ground-state energy of N (noninteracting) electrons in this box. For each energy level associated with a triplet of integers $(1, 1, 1)$, $(2, 1, 1)$, $(1, 2, 1)$, $(1, 1, 2)$, $(3, 1, 1)$, $\dots$, $(2, 2, 2)$, $\dots$, two electrons can be accommodated.

EXAMPLE 13-1

What is the lowest energy of a set of 24 electrons in a cubical box? What would it be in the absence of the exclusion principle?

SOLUTION The energy in units of $\frac{\hbar^2 \pi^2}{2mL^2}$ is $n_1^2 + n_2^2 + n_3^2$ with $(n_1, n_2, n_3) = 1, 2, 3, \dots$. The energies and the number of electrons for each energy, corresponding to the integers are, in order,

$(1, 1, 1) \rightarrow 3$	2 electrons
$(2, 1, 1), (1, 2, 1), (1, 1, 2) \rightarrow 6$	$3 \times 2 = 6$ electrons
$(2, 2, 1), (2, 1, 2), (1, 2, 2) \rightarrow 9$	$3 \times 2 = 6$ electrons
$(3, 1, 1), (1, 3, 1), (1, 1, 3) \rightarrow 11$	$3 \times 2 = 6$ electrons
$(2, 2, 2) \rightarrow 12$	2 electrons
$(3, 2, 1), (3, 1, 2), (2, 3, 1), (2, 1, 3), (1, 2, 3), (1, 3, 2) \rightarrow 14$	12 electrons

We see that energy levels 3, 6, 9, 11, 12 can hold 22 electrons, so that the remaining 2 electrons are in the energy 14 levels. Thus the total energy in our units is $(2 \times 3) + (6 \times 6) + (6 \times 9) + (6 \times 11) + (2 \times 12) + (2 \times 14) = 214$. In the absence of the exclusion principle, all of the electrons would go into the lowest energy state, so that the total energy would be $24 \times 3 = 72$.

A simpler way to deal with the problem of N electrons is to rephrase the question: *How many triplets of integers (n_1, n_2, n_3) are there such that*

$$E = \frac{\hbar^2 \pi^2}{2mL^2} (n_1^2 + n_2^2 + n_3^2) \quad (13-54)$$

is less than some number E_F ?

Each triplet of integers formed lies on a lattice in three-dimensional space, and if there are very many of them, then it is a very good approximation to say that they must lie inside a sphere of radius R given by

$$n_1^2 + n_2^2 + n_3^2 = R^2 = \frac{2mE_F}{\hbar^2 \pi^2} L^2 \quad (13-58)$$

That number is given by the volume of the octant of the sphere for which all the n_i are positive (see Fig. 13-1). Thus the number of lattice points is

$$\frac{1}{8} \cdot \frac{4\pi}{3} R^3 = \frac{1}{8} \frac{4\pi}{3} \left(\frac{2mE_F}{\hbar^2 \pi^2} L^2 \right)^{3/2} \quad (13-59)$$

and hence the number of electrons with energy less than the energy E_F is twice that; that is,

$$N = \frac{\pi}{3} L^3 \left(\frac{2mE_F}{\hbar^2 \pi^2} \right)^{3/2} \quad (13-60)$$

The number of electrons is proportional to the volume of the box L^3 , which is to be expected. In terms of the density of electrons,

$$n = \frac{N}{L^3} \quad (13-61)$$

we have

$$E_F = \frac{\hbar^2 \pi^2}{2m} \left(\frac{3n}{\pi} \right)^{2/3} \quad (13-62)$$

The energy E_F is the energy of the most energetic electron in the ground state of a noninteracting electron gas with density n . It is known as the *Fermi energy* and the labeling was chosen accordingly.

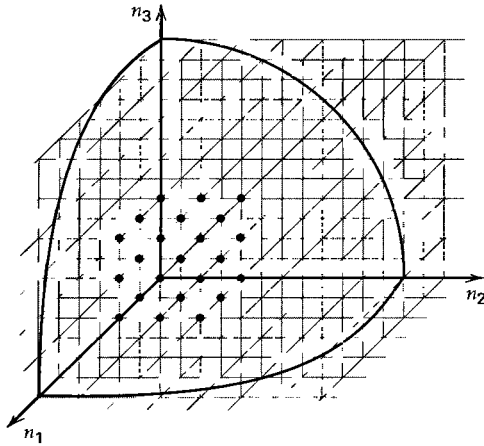


Figure 13-1 Counting of states for an independent particle system.

To calculate the total energy, the number of lattice points may be written as

$$\frac{1}{8} \int_{|\mathbf{n}| \leq R} d^3 \mathbf{n} \quad (13-63)$$

The factor $1/8$ comes from our restriction to positive integers in (13-54); in the integration (13-63) this restriction is removed and must be compensated for by the factor in front. This energy must be doubled, since there are two electrons with the same energy at this lattice point. Hence the total energy is

$$\begin{aligned} E_{\text{tot}} &= \frac{\hbar^2 \pi^2}{mL^2} \frac{1}{8} \int \mathbf{n}^2 d^3 \mathbf{n} \\ &= \frac{\hbar^2 \pi^2}{8mL^2} 4\pi \int_0^R n^4 dn \\ &= \frac{\hbar^2 \pi^3}{10mL^2} R^5 \end{aligned} \quad (13-64)$$

Since R is related to the number of electrons by

$$N = 2 \cdot \frac{1}{8} \cdot \frac{4\pi}{3} R^3 \quad (13-65)$$

we finally get

$$E_{\text{tot}} = \frac{\hbar^2 \pi^3}{10mL^2} \left(\frac{3N}{\pi} \right)^{5/3} \quad (13-66)$$

If we write this in terms of $n = N/L^3$, we get

$$E_{\text{tot}} = \frac{\hbar^2 \pi^3}{10m} \left(\frac{3n}{\pi} \right)^{5/3} L^3 \quad (13-67)$$

The consequences of the Pauli exclusion principle are quite staggering. We will discuss a few of them after the following observations:

- (a) The wave number, defined by $E = \hbar^2 k^2 / 2m$ at the top of the “Fermi sea” is given by

$$k_F = (3\pi^2 n)^{1/3} \quad (13-68)$$

Since $k = 2\pi/\lambda$, we get for the de Broglie wavelength

$$\lambda = 2.03 n^{-1/3} \quad (13-69)$$

Since $n^{-1/3}$ is approximately the interparticle spacing d , we can state our result in the easily memorized form

$$d = \frac{\lambda_F}{2} \quad (13-70)$$

This is a better than expected mnemonic device. Since the exclusion principle forbids two electrons with identical quantum numbers to be on top of each other, this means that they must be at least a half-wave apart.

- (b) If we keep a fixed number of electrons, then (13-66) written in terms of the volume V containing them takes the form

$$E_{\text{tot}} = \frac{\hbar^2 \pi^3}{10m} \left(\frac{3N}{\pi} \right)^{5/3} V^{-2/3} \quad (13-71)$$

This result is actually independent of the shape of the volume, if N is very large. We used a cube to do the computations, since that is the simplest way to do the calculation.

13-6 APPLICATIONS

Degeneracy Pressure and Bulk Modulus

If the electron gas is compressed, the electrons are pushed closer to each other, and this decreases the de Broglie wavelength and, equivalently, increases the kinetic energy. Thus the compression is resisted, and the pressure resisting the compression is called the *degeneracy pressure*. It is given by

$$p_{\text{deg}} = -\frac{\partial E_{\text{tot}}}{\partial V} = \frac{\hbar^2 \pi^3}{15m} \left(\frac{3n}{\pi} \right)^{5/3} \quad (13-72)$$

The bulk modulus B of a material, the reciprocal of the *compressibility*, is defined by

$$B = V \frac{\partial p}{\partial V} \quad (13-73)$$

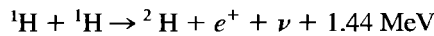
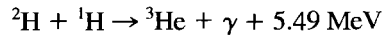
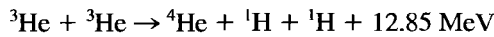
and if we use the degeneracy pressure for p , we find that $B = 5p_{\text{deg}}/3$, so that

$$B = \frac{\hbar^2 \pi^3}{9m} \left(\frac{3n}{\pi} \right)^{5/3} \quad (13-74)$$

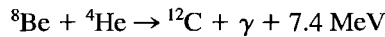
The use of a degenerate electron gas model for a metal gives the correct order of magnitude for the bulk modulus B . For example, for copper we have $n_e = 8.47 \times 10^{28}$ electrons/m³, so that $B = 6.4 \times 10^{10}$ N/m². The experimental value is 14×10^{10} N/m².

Astrophysical Application

The resistance to compression that originates in the Pauli exclusion principle plays an important role in stellar evolution. Stars “burn” by undergoing a succession of nuclear reactions. Hydrogen converts to helium by the series of reactions



When all of the hydrogen is converted to helium, the burning stops. Gravitational contraction compresses the helium until it begins to burn in the reaction



The variety of nuclear processes gets larger, and the process of synthesis of nuclei is now well understood. At some point, when the star consists primarily of iron, silicon, and neighboring elements, the burning stops. The material then resumes gravitational contraction, and the only barrier to total gravitational collapse is the effect of degeneracy pressure.

The gravitational pressure is easily calculated if we assume that the density of material ρ is independent of radius, and that the shape of the star is spherical. The potential energy of the material in a shell lying between radius r and $r + dr$ is

$$dV_g = -G \frac{(4\pi\rho r^3/3)(4\pi\rho r^2 dr)}{r} = -\frac{(4\pi)^2 G \rho^2}{3} r^4 dr \quad (13-75)$$

so that the potential energy of the material contained in a sphere of radius R is

$$V_g = -\frac{(4\pi)^2 G \rho^2}{3} \int_0^R r^4 dr = -\frac{(4\pi)^2}{15} G \rho^2 R^5 \quad (13-76)$$

We also have the connection between ρ , R , and the stellar mass M . The star consists of N nucleons (in the form of iron, silicon, and so on) each of mass m_n , so that

$$\frac{4\pi}{3} \rho R^3 = M = (Nm_n) \quad (13-77)$$

After a little algebra we get an expression for the gravitational potential energy in terms of the volume of the star V :

$$V_g = -\frac{3}{5} \left(\frac{4\pi}{3} \right)^{1/3} G (Nm_n)^2 V^{-1/3} \quad (13-78)$$

The gravitational pressure is

$$p_g = -\frac{\partial V_g}{\partial V} = -\frac{1}{5} \left(\frac{4\pi}{3} \right)^{1/3} G (Nm_n)^2 V^{-4/3} \quad (13-79)$$

This is opposed by the degeneracy pressure, which, according to (13-72) is

$$p_{\text{deg}} = \frac{\hbar^2 \pi^3}{15m_e} \left(\frac{3n}{\pi} \right)^{5/3} = \frac{\hbar^2 \pi^3}{15m_e} \left(\frac{3N_e}{\pi} \right)^{5/3} V^{-5/3} \quad (13-80)$$

where N_e is the number of electrons; it is equal to the number of protons in the star. Assuming equal numbers of protons and neutrons, we have $N_e = N/2$.

The two pressures balance, for a given value of N , when

$$\frac{1}{5} \left(\frac{4\pi}{3} \right)^{1/3} G (Nm_n)^2 V^{-4/3} = \frac{\hbar^2 \pi^3}{15m_e} \left(\frac{3N_e}{\pi} \right)^{5/3} V^{-5/3}$$

that is, when the radius of the star is R^* ,

$$R^* = \left(\frac{3}{4\pi} \right)^{1/3} V^{1/3} = \left(\frac{81\pi^2}{128} \right)^{1/3} \frac{\hbar^2}{Gm_e m_n^2} N^{-1/3} \quad (13-81)$$

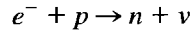
For a star of one solar mass,

$$N = \frac{2 \times 10^{30} \text{ kg}}{1.67 \times 10^{-27} \text{ kg}} = 1.2 \times 10^{57}$$

and the radius of the degenerate star is $R^* \approx 1.1 \times 10^4$ km. The radius of a nondegenerate star, the sun, is $\approx 7 \times 10^8$ km!

If the mass is somewhat larger than a solar mass, the average energy of the electrons increases. When the electrons acquire relativistic energies, our expression for the degeneracy pressure changes drastically. In effect, the electron energy is no longer $p^2/2m_e$, but pc . It can be shown (see Problem 12) that in this domain the degeneracy pressure also scales

as $V^{-4/3}$ and for a sufficiently large value of N , the gravitational pressure overcomes the degeneracy pressure. As a consequence of this large net pressure, the reaction



takes place. The neutrinos escape, since matter, even degenerate matter, is transparent to them, and we are left with a *neutron star*. The degeneracy pressure of the neutrons, which are also fermions, and thus also obey the exclusion principle, can be calculated in the same way as the electron pressure, except that N_e is replaced by N and m_e by m_n . We now obtain

$$R_n^* = \left(\frac{81\pi^2}{16} \right)^{1/3} \frac{\hbar^2}{Gm_n^3} N^{-1/3} \quad (13-82)$$

For a star with two solar masses, we end up with $R_n^* \approx 10$ km! If the mass (equivalently N) is so large that the neutrons become relativistic, then there is no counterbalance to the huge gravitational pressure, and a *black hole* forms.

PROBLEMS

1. What is the reduced mass of an electron–proton system? How does it differ from the reduced mass of an electron–deuteron system? What is the reduced mass of a system of two identical articles?
2. Prove that the exchange operator P_{12} is hermitian.
3. Consider two noninteracting electrons in an infinite potential well. What is the ground-state wave function if the two electrons are in the *same* spin state?
4. Consider N noninteracting electrons in a one-dimensional infinite well of width b . What is the lowest value of the total energy for large N ? [*Hint*: Remember that there are two electrons per energy level; for large N it does not matter whether the last level is filled with one or two electrons, and finally

$$\sum_{n=1}^k n^2 = \frac{k(k+1)(2k+1)}{6} \approx \int_1^k n^2 dn = \frac{k^3}{3} \Big]$$

5. Consider two electrons in the same spin state, interacting with a potential

$$\begin{aligned} V(|x_1 - x_2|) &= -V_0 & |x_1 - x_2| \leq a \\ &= 0 & \text{elsewhere} \end{aligned}$$

What is the lowest energy of the two-electron state, assuming that the total momentum of the two electrons is zero? Assume that the potential is deep enough for more than one bound state. (*Hint*: Reduce the problem to a one-particle equation with reduced mass. Remember that when the electrons are in the same spin state, the spatial wave function must change sign when $x_1 \leftrightarrow x_2$.)

6. Consider two identical particles described by the energy operator

$$H = H(p_1, x_1) + H(p_2, x_2)$$

where

$$H(p, x) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

Separate out the center of mass motion, and obtain the energy spectrum for this system. Show that it agrees with that obtained by solving

$$H\psi(x_1, x_2) = E\psi(x_1, x_2)$$

with

$$\psi(x_1, x_2) = u_1(x_1)u_2(x_2)$$

Discuss the degeneracy of the energy spectrum.

7. Consider two electrons described by the Hamiltonian

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + V(x_1) + V(x_2)$$

where $V(x) = \infty$ for $x < 0$ and for $x > a$; $V(x) = 0$ for $0 < x < a$. Assume that the electrons are in the same spin state; that is, $\sigma_1 = \sigma_2$.

- (a) What is the lowest energy of the two-electron state?
 - (b) What is the energy eigenfunction for this ground state?
 - (c) What is the energy and the wave function of the first excited state, still keeping $\sigma_1 = \sigma_2$?
8. Consider a two-electron system, in which $\sigma_1 = \sigma_2$, so that we do not need to consider spin. Suppose the electrons are in gaussian wave packets about $x = a$ and $x = -a$, so that their wave functions are $\sqrt{\pi/\mu}e^{-\mu^2(x-a)^2/2}$ and $\sqrt{\pi/\mu}e^{-\mu^2(x+a)^2/2}$, respectively. Construct a properly normalized two-electron wave function. Suppose $1/\mu = 0.5 \text{ \AA}$. Estimate for what values of a one can ignore the effects of the Pauli principle to an accuracy of 1 part in 1000.
9. Use the wave function for the two-electron system in Problem 8 to calculate the probability that the separation between the two electrons is in the range $(x, x + dx)$. Also show that the expectation value of the center of mass of the two-electron system is $\langle (x_1 + x_2)/2 \rangle = 0$. (*Hint:* Write x_1 and x_2 in terms of the center-of-mass variable $X = (x_1 + x_2)/2$ and the separation $x = x_1 - x_2$, and express the wave function in terms of the new variables.)
10. Plot the probability density obtained in Problem 9 as a function of x , for the two cases (i) $1/a = \mu/2$, (ii) $1/a = 2\mu$. Discuss the physics behind your results.
11. Suppose the electrons were replaced by bosons in Problems 8–10. List the changes in the formulas, plot the probability density as a function of x for the two separations, and explain the difference between the fermion and boson case for the probability density.
12. Repeat the calculation of the Fermi energy of a gas of fermions by assuming that the fermions are massless, so that the energy-momentum relation is $E = pc$.
13. Calculate the degeneracy of states in a cubic box of volume L^3 as a function of E ; that is, calculate the number of states in the interval $(E, E + dE)$, and use this to obtain the density of states of an electron gas, keeping in mind that there are two electrons per energy state. (*Hint:* How many $\{n_1, n_2, n_3\}$ are there for which $\sum_i n_i^2 = 2mEL^2/\hbar^2\pi^2$?)
14. Calculate the energy density of a gas of photons in a cubic box of volume L^3 , keeping in mind that there are two photons (two states of polarization) per energy state.
15. Calculate the energy spectrum of a gas of photons in a box of sides a, a, L , and $a \ll L$.
16. Calculate the energy spectrum for a free electron in a box of sides a, a, L , with $a \ll L$. Discuss the spacings when $a = 1 \text{ nm}$ and $L = 1 \text{ }\mu\text{m}$.
17. A nucleus consists of N neutrons and Z protons, with $N + Z = A$. If the radius of the nucleus is given by $R = r_0 A^{1/3}$ with $r_0 = 1.1 \text{ fm}$ ($1 \text{ fm} = 10^{-15} \text{ m}$) and if the neutron and proton masses are treated as equal, ($1.7 \times 10^{-27} \text{ kg}$), write an expression for the Fermi energy of the proton “gas” and the neutron “gas,” assuming that the protons and neutrons move as free particles. What are the Fermi energies if $N = 126$ and $Z = 82$?

Supplement 13-A

Conservation of Total Momentum

In our discussion of angular momentum in Chapter 8 we found that the assumption of invariance of the Hamiltonian under rotations led to the appearance of a new constant of motion, *the angular momentum*. In this supplement we show that the assumption of *invariance under spatial displacement* leads to the existence of a constant of the motion, *the momentum*. The requirement that the system be unchanged under the transformation

$$x_i \rightarrow x_i + a \quad (13A-1)$$

does not change the kinetic energy, since a is independent of time. The potential energy will change, *unless* it has the form

$$V(x_1, x_2, x_3, \dots, x_N) = V(x_1 - x_2, x_1 - x_3, \dots, x_{N-1} - x_N) \quad (13A-2)$$

In classical mechanics, the absence of external forces leads to momentum conservation. This follows from the equations of motion,

$$m_i \frac{d^2 x_i}{dt^2} = -\frac{\partial}{\partial x_i} V(x_1 - x_2, x_1 - x_3, \dots, x_{N-1} - x_N) \quad (13A-3)$$

a consequence of which is that

$$\begin{aligned} \frac{d}{dt} \sum_i m_i \frac{dx_i}{dt} &= -\sum_i \frac{\partial}{\partial x_i} V(x_1 - x_2, \dots, x_{N-1} - x_N) \\ &= 0 \end{aligned} \quad (13A-4)$$

The reason for the vanishing of the right side of the preceding equation is that for every argument in V , there are equal and opposite contributions that come from $\sum_i \partial/\partial x_i$ acting on it. For example, with $u = x_1 - x_2$, $v = x_1 - x_3$, $w = x_2 - x_3$,

$$\left(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3} \right) V(u, v, w) = \frac{\partial V}{\partial u} + \frac{\partial V}{\partial v} - \frac{\partial V}{\partial u} + \frac{\partial V}{\partial w} - \frac{\partial V}{\partial v} - \frac{\partial V}{\partial w} = 0$$

Hence

$$P = \sum_i m_i \frac{dx_i}{dt} \quad (13A-5)$$

is a constant of the motion.

In quantum mechanics the same conclusion holds. We shall demonstrate it by using the invariance of the Hamiltonian under the transformation (13A-1). The invariance implies that both

$$Hu_E(x_1, x_2, \dots, x_N) = Eu_E(x_1, x_2, \dots, x_N) \quad (13A-6)$$

and

$$Hu_E(x_1 + a, x_2 + a, \dots, x_N + a) = Eu_E(x_1 + a, x_2 + a, \dots, x_N + a) \quad (13A-7)$$

hold. Let us take a infinitesimal, so that terms of $0(a^2)$ can be neglected. Then

$$\begin{aligned} u(x_1 + a, \dots, x_N + a) &\simeq u(x_1, \dots, x_N) + a \frac{\partial}{\partial x_1} u(x_1, \dots, x_N) \\ &\quad + a \frac{\partial}{\partial x_2} u(x_1, \dots, x_N) + \dots \\ &\simeq u(x_1, \dots, x_N) + a \sum_i \frac{\partial}{\partial x_i} u(x_1, \dots, x_N) \end{aligned} \quad (13A-8)$$

and hence

$$\begin{aligned} aH \left(\sum_{i=1}^N \frac{\partial}{\partial x_i} \right) u_E(x_1, \dots, x_N) &= aE \left(\sum_{i=1}^N \frac{\partial}{\partial x_i} \right) u_E(x_1, \dots, x_N) \\ &= a \left(\sum_{i=1}^N \frac{\partial}{\partial x_i} \right) E u_E(x_1, \dots, x_N) \\ &= a \left(\sum_{i=1}^N \frac{\partial}{\partial x_i} \right) H u_E(x_1, \dots, x_N) \end{aligned} \quad (13A-9)$$

If we now define

$$P = \frac{\hbar}{i} \sum_{i=1}^N \frac{\partial}{\partial x_i} \equiv \sum_{i=1}^N p_i \quad (13A-10)$$

so that P is the total momentum operator, we see that we have demonstrated that

$$(HP - PH) u_E(x_1, \dots, x_N) = 0 \quad (13A-11)$$

Since the energy eigenstates for N -particles form a complete set of states, in the sense that any function of $x_1, x_2, \dots, x_N$ can be expanded in terms of all the $u_E(x_1, \dots, x_N)$, the preceding equation can be translated into

$$[H, P]\psi(x_1, \dots, x_N) = 0 \quad (13A-12)$$

for all $\psi(x_1, \dots, x_N)$ —that is, into the operator relation

$$[H, P] = 0 \quad (13A-13)$$

This, however, implies that P , the total momentum of the system, is a *constant of the motion*. This is a very deep consequence of what is really a statement about the nature of space. The statement that there is no origin—that is, that the laws of physics are invariant under displacement by a fixed distance—leads to a conservation law. In relativistic quantum mechanics there are no potentials of the form that we consider here; nevertheless the invariance principle, as stated earlier, still leads to a conserved total momentum.

Chapter 14

About Atoms and Molecules

Hydrogen and hydrogenlike atoms have only one electron. The force on the electron is simple: It is just the attractive Coulomb force due to the charge of the nucleus. When more electrons are involved, the potential energy involves more than the interaction with the nucleus; the repulsion between the electrons comes into play. All the complications that arise from this additional interaction and *the effect of the Pauli exclusion principle* show up in the simplest two-electron system—the helium atom. Our discussion of this atom will serve as a model for the discussion of more complex atoms.

The simplest molecule is H_2^+ , in which there is only one electron but two protons. Our discussion of this system will provide us with some guidance on how to treat more complicated diatomic molecules. Much of this material is in a supplement, but we do discuss the lowest excitations in diatomic molecules, the rotational spectra, in the text.

14-1 THE HELIUM ATOM WITHOUT ELECTRON-ELECTRON REPULSION

The helium atom consists of a nucleus of charge $Z = 2$ and two electrons, which we label 1 and 2. Each electron is attracted to the nucleus, and the two electrons repel each other. We assume, and this turns out to be correct, that no forces other than the electromagnetic ones (Coulomb to a very good approximation), are necessary to describe the dynamics of the helium atom with the help of quantum mechanics.

If the nucleus is placed at the origin, and if the electron coordinates are labeled $\mathbf{r}_1$ and $\mathbf{r}_2$, then the Hamiltonian for the atom is (Fig. 14-1)

$$H = \frac{\mathbf{p}_1^2}{2m_e} + \frac{\mathbf{p}_2^2}{2m_e} - \frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{Ze^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} \quad (14-1)$$

We shall ignore the small effects concerned with the motion of the nucleus,¹ relativistic effects, spin-orbit couplings, and the effect of the current caused by the motion of one electron on the other electron. The Hamiltonian just given can be written as

$$H = H^{(1)} + H^{(2)} + V_{12} \quad (14-2)$$

where

$$H^{(i)} = \frac{\mathbf{p}_i^2}{2m_e} - \frac{Ze^2}{4\pi\epsilon_0 r_i} \quad (14-3)$$

¹The reduced mass effect takes a somewhat different form, since one is trying to convert a three-body problem into an effective two-body problem. This is worked out in D. Park, *Introduction to the Quantum Theory* (3rd ed.), McGraw-Hill, New York, 1992.

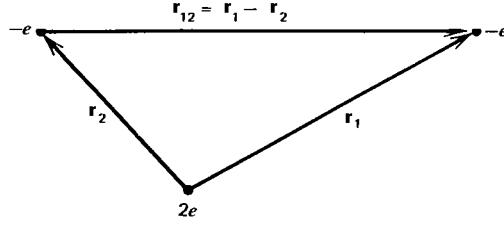


Figure 14-1 Coordinates used in the description of the helium atom.

and

$$V_{12} = \frac{e^2}{4\pi\epsilon_0|\mathbf{r}_1 - \mathbf{r}_2|} \quad (14-4)$$

We shall work with the nuclear charge Z and set $Z = 2$ later. Our work on the hydrogen atom provides us with a complete set of eigenfunctions for $H^{(1)}$ and $H^{(2)}$. Thus, if we were to ignore V_{12} in the total Hamiltonian, we would have a solution to the eigenvalue problem for the two-electron system. The eigenfunctions would be

$$u(\mathbf{r}_1, \mathbf{r}_2) = \phi_{n_1 l_1 m_1}(\mathbf{r}_1) \phi_{n_2 l_2 m_2}(\mathbf{r}_2) \quad (14-5)$$

for the equation

$$[H^{(1)} + H^{(2)}] u(\mathbf{r}_1, \mathbf{r}_2) = E u(\mathbf{r}_1, \mathbf{r}_2) \quad (14-6)$$

and the energy would be given by (Fig. 14.2a)

$$E = E_{n_1} + E_{n_2} \quad (14-7)$$

where $E_n = -(mc^2/2)(Z\alpha)^2/n^2$. Thus in the idealized model in which the two electrons ignore each other, the lowest energy is

$$E = -2E_1 = -mc^2(2\alpha)^2 = -108.8 \text{ eV} \quad (14-8)$$

Note that this is $2 \times Z^2 = 8$ times the hydrogen energy of -13.6 eV .

The first excited state is one in which one electron is in its ground state, $n = 1$, and the second electron is raised to the first excited $n = 2$ state. Then

$$E = E_1 + E_2 = -68.0 \text{ eV} \quad (14-9)$$

The ionization energy—that is, the energy required to remove one electron from the ground state to infinity—is

$$E_{\text{ioniz}} = (E_1 + E_\infty) - 2E_1 = 54.4 \text{ eV} \quad (14-10)$$

and, interestingly enough, the onset of the continuum lies *lower* than the excited state for which both electrons are in the $n = 2$ state. The energy of the latter state is

$$E = 2E_2 = -27.2 \text{ eV} \quad (14-11)$$

and it brings up a new phenomenon: the existence of a discrete state in the continuum for the Hamiltonian $H^{(1)} + H^{(2)}$. We shall briefly discuss the implications of this later in the chapter, in the section on *Autoionization*.

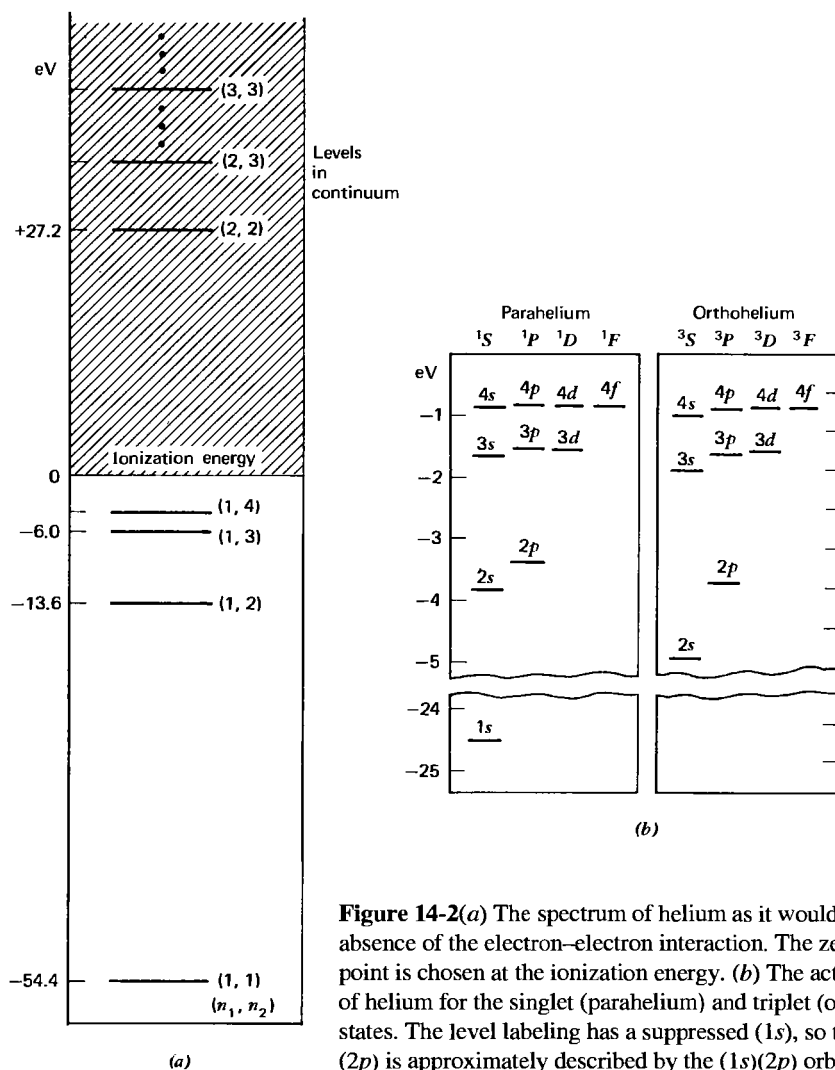


Figure 14-2(a) The spectrum of helium as it would look in the absence of the electron-electron interaction. The zero energy point is chosen at the ionization energy. **(b)** The actual spectrum of helium for the singlet (parahelium) and triplet (orthohelium) states. The level labeling has a suppressed $(1s)$, so that the level $(2p)$ is approximately described by the $(1s)(2p)$ orbital.

Effects of the Exclusion Principle

Since the two electrons are *identical fermions* we must make the total wave function anti-symmetric under the interchange of space and spin coordinates of the electrons. Thus a proper description of the ground state of this idealized model is

$$u_0(\mathbf{r}_1, \mathbf{r}_2) = \phi_{100}(\mathbf{r}_1) \phi_{100}(\mathbf{r}_2) X_{\text{singlet}} \quad (14-12)$$

The spatial part of the wave function is necessarily symmetric, and that is why the state must be a spin singlet state

$$X_{\text{singlet}} = \frac{1}{\sqrt{2}} (\chi_+^{(1)} \chi_-^{(2)} - \chi_-^{(1)} \chi_+^{(2)}) \quad (14-13)$$

For the first excited state, we have two possibilities, which, for $V = 0$, are degenerate in energy. These are

$$u_1^{(s)} = \frac{1}{\sqrt{2}} [\phi_{100}(\mathbf{r}_1) \phi_{2lm}(\mathbf{r}_2) + \phi_{2lm}(\mathbf{r}_1) \phi_{100}(\mathbf{r}_2)] X_{\text{singlet}} \quad (14-14)$$

and the space-antisymmetric, spin symmetric

$$u_1^{(t)} = \frac{1}{\sqrt{2}} [\phi_{100}(\mathbf{r}_1) \phi_{2lm}(\mathbf{r}_2) - \phi_{2lm}(\mathbf{r}_1) \phi_{100}(\mathbf{r}_2)] X_{\text{triplet}} \quad (14-15)$$

where

$$X_{\text{triplet}} = \begin{cases} \chi_+^{(1)} \chi_+^{(2)} \\ \frac{1}{\sqrt{2}} (\chi_+^{(1)} \chi_-^{(2)} + \chi_-^{(1)} \chi_+^{(2)}) \\ \chi_-^{(1)} \chi_-^{(2)} \end{cases} \quad (14-16)$$

is orthogonal to X_{singlet} .

14-2 EFFECTS OF ELECTRON-ELECTRON REPULSION

The presence of V , the electron-electron Coulomb repulsion, may, in first approximation, be treated as a perturbation. Let us first compute the energy shift of the ground state to first order in V . We need to evaluate

$$\Delta E = \iint d^3r_1 d^3r_2 u_0^*(\mathbf{r}_1, \mathbf{r}_2) \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} u_0(\mathbf{r}_1, \mathbf{r}_2) \quad (14-17)$$

We start with a crude estimate of the effect of this perturbation. Each electron is approximately a distance a_0/Z from the nucleus. The energy is lowest when the electrons are as far away from each other as possible. This means that the effective separation between the electrons will be fa_0/Z , where f is some number, no larger than 2. The repulsion energy is then

$$\Delta E \approx \frac{e^2}{4\pi\epsilon_0 fa_0} \frac{Z}{f} = \frac{Z}{f} (27.2 \text{ eV}) \quad (14-18)$$

where in the last step we have used the fact that

$$\frac{e^2}{4\pi\epsilon_0} \frac{1}{2a_0} = 13.6 \text{ eV}$$

Since the perturbation does not involve spin, we need only consider

$$\Delta E = \iint d^3r_1 d^3r_2 |\phi_{100}(\mathbf{r}_1)|^2 \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} |\phi_{100}(\mathbf{r}_2)|^2 \quad (14-19)$$

The integral has a simple interpretation: Since $|\phi_{100}(\mathbf{r}_1)|^2$ is the probability density of finding electron 1 at $\mathbf{r}_1$ we may interpret $e|\phi_{100}(\mathbf{r}_1)|^2$ as the charge density for electron 1. Hence

$$U(\mathbf{r}_2) = - \int d^3r_1 \frac{e|\phi_{100}(\mathbf{r}_1)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \quad (14-20)$$

is the potential at $\mathbf{r}_2$ due to the charge distribution of electron 1, and

$$\Delta E = -\frac{1}{4\pi\epsilon_0} \int d^3r_2 e |\phi_{100}(\mathbf{r}_2)|^2 U(\mathbf{r}_2)$$

is therefore the electrostatic energy of interaction of electron 2 with that potential. The integral can be worked out. With

$$\phi_{100}(\mathbf{r}) = \frac{2}{\sqrt{4\pi}} \left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0} \quad (14-21)$$

we have

$$\Delta E = \left(\frac{1}{\pi} \left(\frac{Z}{a_0}\right)^3\right)^2 \frac{e^2}{4\pi\epsilon_0} \int_0^\infty r_1^2 dr_1 e^{-2Zr_1/a_0} \int_0^\infty r_2^2 dr_2 e^{-2Zr_2/a_0} \int d\Omega_1 \int d\Omega_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \quad (14-22)$$

We shall use the fact that

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{1}{(r_1^2 + r_2^2 - 2r_1r_2 \cos \theta)^{1/2}} \quad (14-23)$$

where θ is the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$. We may proceed in one of two ways.

- (a) Most directly, we choose the direction of $\mathbf{r}_1$ as the z -axis for the $d\Omega_2$ integration, and get

$$\begin{aligned} \int d\Omega_2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} &= \int_0^{2\pi} d\phi \int_{-1}^1 d(\cos \theta) \frac{1}{(r_1^2 + r_2^2 - 2r_1r_2 \cos \theta)^{1/2}} \\ &= -2\pi \frac{1}{r_1r_2} [(r_1^2 + r_2^2 - 2r_1r_2 \cos \theta)^{1/2}]_{\cos \theta = -1}^{\cos \theta = +1} \\ &= \frac{2\pi}{r_1r_2} (r_1 + r_2 - |\mathbf{r}_1 - \mathbf{r}_2|) \end{aligned} \quad (14-24)$$

The integration over $d\Omega_1$ is trivial since nothing depends on that direction, and it gives 4π . This leaves us with

$$8 \frac{e^2}{4\pi\epsilon_0} \left(\frac{Z}{a_0}\right)^6 \int_0^\infty r_1 dr_1 e^{-2Zr_1/a_0} \int_0^\infty r_2 dr_2 e^{-2Zr_2/a_0} (r_1 + r_2 - |\mathbf{r}_1 - \mathbf{r}_2|) \quad (14-25)$$

- (b) A very useful expansion, necessary when there is additional angular dependence in the numerator, is the following. For $r_1 > r_2$,

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{1}{(r_1^2 + r_2^2 - 2r_1r_2 \cos \theta)^{1/2}} = \frac{1}{r_1} \sum_{L=0}^\infty \left(\frac{r_2}{r_1}\right)^L P_L(\cos \theta) \quad (14-26)$$

with the roles of r_1 and r_2 reversed when $r_2 > r_1$.

We next evaluate (14-25), which gives us

$$\Delta E = \frac{8e^2}{4\pi\epsilon_0} \left(\frac{Z}{a_0}\right)^6 \int_0^\infty r_1 dr_1 e^{-2Zr_1/a_0} \left\{ 2 \int_0^{r_1} r_2^2 dr_2 e^{-2Zr_2/a_0} + 2r_1 \int_{r_1}^\infty r_2 dr_2 e^{-2Zr_2/a_0} \right\}$$

The integrals are standard and yield the answer²

$$\Delta E = \frac{5}{8} \frac{Ze^2}{4\pi\epsilon_0 a_0} = \frac{5}{4} Z \left(\frac{1}{2} mc^2 \alpha^2 \right) \quad (14-27)$$

This is a positive contribution, since it arises from a repulsive force, and its magnitude for $Z = 2$ is 34 eV. When this is added to the zero-order result of -108.8 eV we obtain, to first order,

$$E \approx -74.8 \text{ eV}$$

When this is compared with

$$E_{\text{exp}} = -78.975 \text{ eV}$$

a sizable discrepancy is seen. Physically, we can attribute this discrepancy to the fact that in our calculation we took no account of “screening”—that is, the effect that the presence of one electron tends to decrease the net charge “seen” by the other electron. Very roughly, if one argues that, for example, electron 1 is half the time “between” electron 2 and the nucleus, then half the time electron 2 sees a charge Z and half the time it sees a charge $Z - 1$; that is, effectively, in the expression

$$E + \Delta E = -\frac{1}{2} mc^2 \alpha^2 \left(2Z^2 - \frac{5}{4} Z \right) \quad (14-28)$$

$(Z - 1/2)$ should be substituted for Z . This does improve agreement, but the crude argument advanced is not sufficient justification of the choice of 50 percent for the probability of effective screening. We will return to this subject later in this chapter, when we discuss the Rayleigh-Ritz variational principle for the ground-state energy.

14-3 EXCLUSION PRINCIPLE AND EXCHANGE INTERACTION

We next consider the first excited state of helium. It will be sufficient to calculate the energy shift with the singlet and triplet $m = 0$ states listed in (14-14) and (14-15), since the shift is caused by a perturbation that commutes with L_z . For such a perturbation, the shift must be independent of the m -value. Again, because of the spin-independence of the perturbing potential, V , we have

$$\begin{aligned} \Delta E_1^{(s,t)} &= \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \int d^3r_1 \int d^3r_2 |\phi_{100}(\mathbf{r}_1)\phi_{200}(\mathbf{r}_2) \pm \phi_{200}(\mathbf{r}_1)\phi_{100}(\mathbf{r}_2)|^2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ &= \frac{e^2}{4\pi\epsilon_0} \int d^3r_1 \int d^3r_2 |\phi_{100}(\mathbf{r}_1)|^2 |\phi_{200}(\mathbf{r}_2)|^2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \\ &\quad \pm \frac{e^2}{4\pi\epsilon_0} \int d^3r_1 \int d^3r_2 \phi_{100}^*(\mathbf{r}_1)\phi_{200}^*(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \phi_{200}(\mathbf{r}_1)\phi_{100}(\mathbf{r}_2) \end{aligned} \quad (14-29)$$

In obtaining this simplified form, we made use of the symmetry of V under $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$.

The energy shift is seen to consist of two terms: The first has the familiar form of an electrostatic interaction between two “electron clouds” distributed according to the wave functions of the two electrons. This term is just a simple generalization of the term that we

²Note that this corresponds to a value of $f = 1.6$ in (14-18).

found for the ground-state energy shift. The second term has no classical interpretation. Its origin lies in the Pauli principle, and its sign depends on whether the state has spin 0 or 1. Thus, because of this *exchange* contribution, the singlet and triplet terms are no longer degenerate. Although we considered $n = 2$ here, we have quite generally

$$\begin{aligned}\Delta E_{n,l}^{(o)} &= J_{nl} - K_{nl} \\ \Delta E_{n,l}^{(s)} &= J_{nl} + K_{nl}\end{aligned}\quad (14-30)$$

The integrals can be evaluated in closed form [it is here that (14-26) becomes useful], but we shall not do this here. The integral J_{nl} is manifestly positive, and it turns out that this is also the case for K_{nl} . For $l = n - 1$ this is obvious: The wave functions appearing in (14-29) have no nodes in that case. That the triplet state should have a lower energy than the singlet state—that is, that

$$J_{nl} - K_{nl} < J_{nl} + K_{nl}$$

or equivalently

$$K_{nl} > 0 \quad (14-31)$$

can be argued on qualitative grounds. For the triplet state the spatial wave function is anti-symmetric, so that the electrons are somewhat constrained to stay away from each other. This tends to reduce the screening effect, so that each electron “sees” more of the nuclear charge, and it also tends to make the repulsion between the electrons less effective than for the spatially symmetric singlet state. An interesting aspect of this result is that, although the perturbing potential (14-4) does not depend on the spins of the electrons, the symmetry of the wave function does make the potential act as if it were spin-dependent. We may write (14-30) in a form that exhibits this. Let the spins of the two electrons be s_1 and s_2 . Then the total spin $S = s_1 + s_2$, and

$$S^2 = s_1^2 + s_2^2 + 2s_1 \cdot s_2 \quad (14-32)$$

If we act with this on triplet and singlet states (14-16) and (14-13) that are also eigenstates of s_1^2 and s_2^2 , we get

$$S(S+1)\hbar^2 = \frac{3}{4}\hbar^2 + \frac{3}{4}\hbar^2 + 2s_1 \cdot s_2 \quad (14-33)$$

that is,

$$2s_1 \cdot s_2 / \hbar^2 = S(S+1) - \frac{3}{2} = \begin{cases} \frac{1}{2} & \text{triplet} \\ -\frac{3}{2} & \text{singlet} \end{cases} \quad (14-34)$$

We may thus write, in terms of the σ 's related to the spins by $s_i = (1/2)\hbar\sigma_i$,

$$\Delta E_{n,l} = J_{n,l} - \frac{1}{2}(1 + \sigma_1 \cdot \sigma_2)K_{nl} \quad (14-35)$$

This result has implications beyond the explanation of some of the details of the spectrum of helium. As was first pointed out by Heisenberg, the *exchange forces* provide a mechanism by means of which spin-dependent effects are of magnitude comparable to those that are independent of spin. Usually, as illustrated in spin-orbit coupling or hyperfine coupling (see Ch. 12) spin-dependent forces have a magnetic origin, and are thus reduced by factors of $O(v^2/c^2) = O(\alpha^2)$ compared to the electrical forces. Such weak forces could not be strong enough to keep the electron spins aligned in a ferromagnet, except at

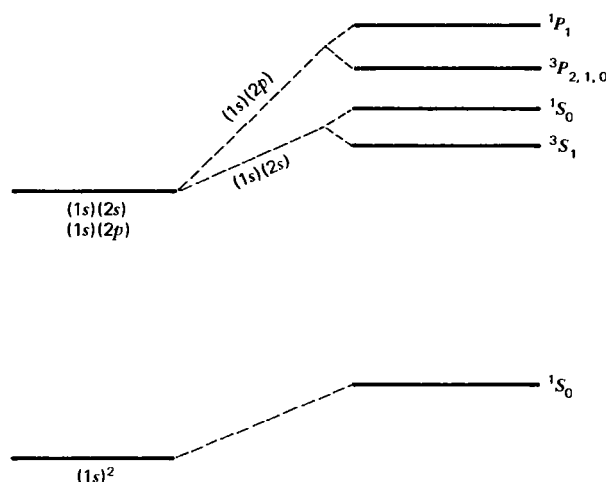


Figure 14-3 Schematic sketch of splitting of the first excited states of helium.

extremely low temperatures. In fact, exchange forces are responsible for the phenomenon of ferromagnetism.

The spectrum of the first few excited states of helium is shown in Fig. 14-3. The notation used for the unperturbed states is that of *orbitals*—that is, the quantum numbers of the unperturbed electrons. Thus both electrons in the ground state are in $n = 1$, $l = 0$ states, and we write this as $(1s, 1s)$, or more briefly $(1s)^2$. It should be understood that when we write $(1s)(2p)$, as for the first excited state, this does not mean that one electron is in one state and the other electron in the other, since we must write totally antisymmetric wave functions for the electrons. Another way of labeling the state is by the $^{2S+1}L_J$ notation, which we use for the perturbed states in the figure. We see that the singlet states lie above the triplet states in a given multiplet. This follows from the symmetry (cf. our argument that $K_{nl} > 0$) and is a special example of one of *Hund's rules*: *Other things being equal, the states of highest spin will have the lowest energy*.³

If we excite helium from the ground state by shining ultraviolet light on it, we find that the *selection rule* $\Delta L = 1$, which we will derive later, implies an excitation to the P states. Furthermore, there is a selection rule $\Delta S = 0$; that is, only transitions singlet $\rightarrow$ singlet and triplet $\rightarrow$ triplet dominate.⁴ Hence the state most strongly excited from the ground state is the $^1P_1$ state. The other levels may also become occupied through other mechanisms—for example, collisional excitation. Once occupied, the radiative transitions to the ground state are very improbable. The ^3P state, which may be populated when atoms in the $^1P_1$ state undergo collisions with other atoms in the gas, can only decay to the $^3S_1$ state, and that state is *metastable*, since it cannot decay to the ground state easily. The fact that there are no transitions, to good approximation, between triplet states and singlet states led, at one time, to the belief that there existed two kinds of helium, orthohelium (triplet) and parahelium (singlet).

³Hund's rules are discussed in more detail in Supplement 14-B [www.wiley.com/college/gesiorowicz].

⁴Selection rules will be discussed in detail in Chapter 17. At this point we merely note that a hint of their existence was seen in our discussion of the $\Delta n = 1$ rule in the old quantum theory Bohr atom. Their origin is angular momentum conservation, together with the property of radiation that at low frequencies photons effectively act as if they had angular momentum 1.

The spectrum of helium that we saw in Fig. 14-2*b* shows that the excited states $(1s)(nl)$ have energies that do not differ very much from those of the hydrogen atom levels. Thus the binding energy of one electron in the atom is 24.6 eV (total binding energy minus binding energy of singly ionized helium = $79.0 - 54.4 = 24.6$ eV), whereas the energy that would be liberated if one electron were to be removed from the $2s$ state is of the order of $4 - 5$ eV, which is comparable to the energy 3.4 eV ($= 13.6/n^2$ eV) for hydrogen. The reason for this effect is that the “outer” electron sees only a unit positive charge, since the “inner” electron in the $(1s)$ orbital tends to shield the nucleus, leaving a net effective charge $\approx Z - 1$. This is not the case for the ground state, since both electrons have access to the nucleus. Thus the ground state lies quite a bit deeper than the hydrogen ground state.

Our discussion shows that the perturbative calculation of the ground-state energy differs from the experimental value by approximately 4 eV. A second-order perturbation calculation is prohibitively difficult. In the days before the three-body problem could easily be calculated with high-powered computers, another method was found that turned out to be very useful for the calculation of the ground-state energy. The method is known as the *Ritz variational principle*, and we shall discuss it later in the chapter. At this point we return to our observation that there exist eigenvalues of $H^{(1)} + H^{(2)}$ that lie above the ionization threshold and that are nevertheless discrete. This topic goes by the name of *autoionization*.

Autoionization

Let us examine the implications of the existence of states that lie above the ionization energy. Examples of such states are, in orbital language, $(2s)^2$, $(2s)(2p)$. Their existence has some dramatic physical consequences. Let us consider, for example, the $(2s)(2p)$ state. If the electrons form a spin singlet state, then this will be a $^1P_1$ state (since $S = 0$, $J = L$, and with $l_1 = 0$, $l_2 = 1$, $L = 1$). As such, it can be excited from the $^1S_0$ ground state by the absorption of radiation, since the selection rules $\Delta L = 1$ and $\Delta S = 0$ are not being violated.

This state, once excited, need not decay back to the ground state ($^1S_0$) or to another state allowed by the selection rules (a $^1D_2$ state, say), because it can go into another *channel*: It can decay into an electron and singly ionized helium, He^+ , with the electron energy determined by energy conservation. This process is described as *autoionization*.

The $(2s)(2p)$ state in the continuum will show up very clearly in the scattering of electrons by He^+ ions. When the electron energy is such that the *compound state* can be formed, a very dramatic peak will occur in the scattering rate. Similarly, in the absorption of radiation by helium, in the vicinity of the energy of the compound state ($e^- + \text{He}^+$), a sharp peak is seen in the absorption (Fig. 14-4). There is absorption at other energies, too, since the process



can occur, but the absorption at energies away from the compound state energy will vary very smoothly with energy. We can describe the state in still another way by calling it a *resonant state*. Since it decays into its constituents $e^- + \text{He}^+$, it does not exist forever. Hence, by the uncertainty relation, $\Delta E \approx \hbar/\Delta t$, it appears that its energy is not precisely defined, which seems to contradict the fact that the $(2s)(2p)$ state does have a well-defined energy. It turns out that if the coupling of the discrete state to the continuum state is taken into account, the state ceases to be discrete, and its energy could lie anywhere in a narrow range about the energy as calculated without the coupling.

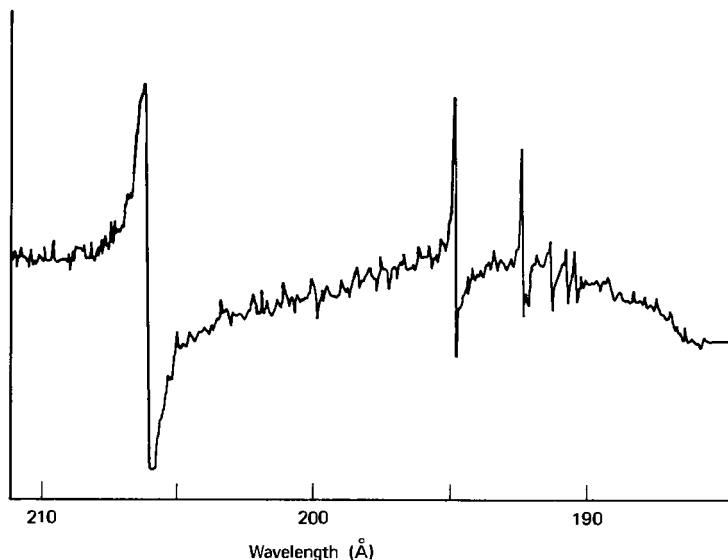


Figure 14-4 Resonance in the helium absorption spectrum above the continuum threshold; the first peak occurs at the energy corresponding to the location of the $(2s)(2p)$ level. (From R. P. Madden and K. Codling, *Phys. Rev. Lett.* **10**, 516 (1963), by permission.)

14-4 THE RITZ VARIATIONAL PRINCIPLE

A very useful calculational tool for the study of ground-state wave functions is the *Ritz variational principle*. The basic idea is the following:

Consider a Hamiltonian H and an *arbitrary* square integrable state vector $|\Psi\rangle$, which we choose to be normalized to unity so that

$$\langle\Psi|\Psi\rangle = 1 \quad (14-36)$$

This function can be expanded in a complete set of eigenstates of H denoted by $|k\rangle$, where $H|k\rangle = E_k|k\rangle$. The expansion has the form

$$|\Psi\rangle = \sum_k C_k |k\rangle \quad (14-37)$$

Now

$$\begin{aligned} \langle\Psi|H|\Psi\rangle &= \sum_{n,m} C_n^* C_m \langle n|H|m\rangle = \sum_{n,m} C_n^* C_m E_m \langle n|m\rangle \\ &= \sum_m |C_m|^2 E_m \end{aligned} \quad (14-38)$$

Since (14-36) implies

$$\sum_m |C_m|^2 = 1 \quad (14-39)$$

and E_0 is the lowest of the eigenvalues (it is the ground state), it follows that

$$\langle\Psi|H|\Psi\rangle \geq E_0 \sum_m |C_m|^2 = E_0 \quad (14-40)$$

We may use this result to calculate an upper bound on E_0 . We choose a $|\Psi\rangle$ that depends on a number of parameters ($\alpha_1, \alpha_2, \dots$), calculate $\langle\Psi|H|\Psi\rangle$, and minimize this with respect to the parameters, making sure that the normalization condition $\langle\Psi|\Psi\rangle = 1$ is maintained. The usefulness of this procedure depends on a good choice of the state $|\Psi\rangle$, a choice that should approximate the true ground state. This procedure then requires some

understanding of the properties of the ground state. We proceed with two illustrations of the usefulness of the variational principle.

The Ground-State Energy of Helium

We choose for the wave function representing $|\Psi\rangle$ a product of hydrogenlike wave function in the (1s) orbital, corresponding to an arbitrary charge Z^* . This is then the variational parameter corresponding to the set $(\alpha_1, \alpha_2, \dots)$. We take

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_{100}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2) \quad (14-41)$$

where

$$\left(\frac{\mathbf{p}^2}{2m_e} - \frac{Z^*e^2}{4\pi\epsilon_0 r} \right) \psi_{100}(\mathbf{r}) = \epsilon \psi_{100}(\mathbf{r}) \quad (14-42)$$

with $\epsilon = -\frac{1}{2} m_e c^2 (Z^* \alpha)^2$, and where the ψ_{100} are normalized. We now need

$$\int d^3r_1 \int d^3r_2 \psi_{100}^*(\mathbf{r}_1) \psi_{100}^*(\mathbf{r}_2) \left\{ \frac{\mathbf{p}_1^2}{2m_e} + \frac{\mathbf{p}_2^2}{2m_e} - \frac{Ze^2}{4\pi\epsilon_0 r_1} - \frac{Ze^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_1 - \mathbf{r}_2|} \right\} \psi_{100}(\mathbf{r}_1) \psi_{100}(\mathbf{r}_2) \quad (14-43)$$

Consider the first and third terms. If the Z were replaced by Z^* , we would get ϵ for the expectation value. If we write the potential energy in the above equation as $(e^2/4\pi\epsilon_0 r_1) Z^* + (e^2/4\pi\epsilon_0 r_1) (Z - Z^*)$, we get for the first and third terms

$$\begin{aligned} \epsilon - \frac{e^2(Z - Z^*)}{4\pi\epsilon_0} \int d^3r_1 \frac{|\psi_{100}(\mathbf{r}_1)|^2}{r_1} &= \epsilon - \frac{e^2(Z - Z^*)}{4\pi\epsilon_0} \left\langle \frac{1}{r_1} \right\rangle_{100} \\ &= \epsilon - \frac{e^2(Z - Z^*)}{4\pi\epsilon_0} \frac{Z^*}{a_0} \end{aligned} \quad (14-44)$$

In the last step we used (8-52) for a hydrogenlike atom with nuclear charge Z^* . An identical factor comes from the second and fourth terms. The fifth term has already been calculated in (14-26). Adding all these up we get

$$\begin{aligned} \langle \Psi | H | \Psi \rangle &= -\frac{1}{2} m_e c^2 \alpha^2 \left(2Z^{*2} + 4Z^*(Z - Z^*) - \frac{5}{4} Z^* \right) \\ &= -\frac{1}{2} m_e c^2 \alpha^2 \left(-2Z^{*2} + 4ZZ^* - \frac{5}{4} Z^* \right) \end{aligned} \quad (14-45)$$

Minimizing this with respect to Z^* yields

$$Z^* = Z - \frac{5}{16} \quad (14-48)$$

which is an improvement on the guess we made earlier ($Z - 1/2$). We thus obtain

$$E_0 \leq -\frac{1}{2} m_e c^2 \alpha^2 \left[2 \left(Z - \frac{5}{16} \right)^2 \right] = -77.38 \text{ eV} \quad (14-47)$$

when we substitute $Z = 2$. This is much better than the first-order perturbation result.

To get better agreement with experiment, the trial wave function $\psi(\mathbf{r}_1, \mathbf{r}_2)$ must include correlations between the electrons that reflect the fact that their interaction has more consequences than just screening. These matters are beyond the scope of this book. The second illustration follows on pp. 229–230.

14-5 THE ATOM WITH Z ELECTRONS⁵

The energy eigenvalue problem for an atom with Z electrons has the form

$$\left(\sum_{i=1}^Z \frac{\mathbf{p}_i^2}{2m_e} - \sum_{i=1}^Z \frac{Ze^2}{4\pi\epsilon_0 r_i} + \frac{e^2}{4\pi\epsilon_0} \frac{1}{2} \sum_{j \neq i}^Z \sum_{i=1}^Z \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) = E \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) \quad (14-48)$$

This is a partial differential equation in $3Z$ dimensions. In the late 1920s and early 1930s a number of physicists developed methods to study the eigenvalues with what were then very primitive computers. In the last 50 years the detailed understanding of atomic structure has become the province of quantum chemists. For the purposes of this book, we shall be satisfied with a qualitative, intuitive understanding of the ground-state wave functions and eigenvalues for atoms with $Z < 40$ or so.

To do so, we shall make use of an approximation, which actually works very well, and which gives us much physical insight into the structure of atoms. It is the *independent electron approximation*. Here one assumes that the effect of the electron–electron repulsion term, in which the potential energy of any one electron at some position depends on the position of each one of the other electrons at the same time, has the net effect of providing an *effective potential* for any one electron.⁶ This effective potential, which acts in addition to the Coulomb attraction to the nucleus of charge Ze , is an average *screening potential*. A great simplification arises if one further assumes that this effective potential is spherically symmetric.

Physically this is not a separate assumption, because the smearing of charges “seen” by a particular electron, caused by the rapid motion of the remaining $Z - 1$ electrons, will also tend to wipe out directional effects. Now for a spherically symmetric potential we know that we can label the energy eigenvalues by the quantum number n (for the energy), l , m , and m_s , representing the orbital angular momentum, the z -component of the angular momentum, and the spin state $\pm 1/2$ of the electron. For the potential we adopt a crude form that describes screening. We take

$$V_{\text{eff}}(r) = \frac{e^2}{4\pi\epsilon_0} \frac{Z(r)}{r} \quad (14-49)$$

$Z(r)$ is the effective charge “seen” by the electron, with $Z(r) \rightarrow Z$ as $r \rightarrow 0$ and $Z(r) \rightarrow 1$ as $r \rightarrow \infty$. Note that each electron experiences this potential. Each electron will therefore have available to it a set of energy levels that are somewhat similar to those of hydrogen. The energy levels will be labeled by the quantum number n , l , m_l , m_s . Since the potential is central, the states for $m_l = l, l - 1, l - 2, \dots, -l$ will be degenerate. Furthermore, with the neglect of any spin-orbit coupling, each of the levels will be twofold degenerate, corresponding to up and down spin states. We do not expect the further de-

⁵A thorough discussion may be found in H. A. Bethe and R. W. Jackiw, *Intermediate Quantum Mechanics*, Perseus Press Cambridge, Mass., 1997.

⁶A formal discussion of this can be found in Supplement 14-A [www.wiley.com/college/gasiorowicz].

generacy of all the states that correspond to a fixed n , since the potential no longer has a pure $1/r$ dependence.

The lowest energy state will be the $n = 1, l = 0$ state. We describe it as the $(1s)$ *orbital*. The next two orbitals $(2s)$ and $(2p)$ will not be degenerate. Quite generally, we expect the $(2p)$ states to have a slightly higher energy than the $(2s)$ states. The reason is that the $(2p)$ wave functions are pushed away from the origin more than those of the $(2s)$ states, and thus they “see” less of the nuclear charge. Similarly, the sequence of states from lower to higher energies for the $n = 3$ states are $(3s)$, $(3p)$, and $(3d)$. We can now take into account the Pauli principle. In any orbital, the number of different (degenerate) states is $(2s + 1)(2l + 1)$. The first factor is always 2, so that for in any (ns) state we can accommodate 2 electrons, in a (np) state $2 \times 3 = 6$ electrons, in a (nd) state $2 \times 5 = 10$ electrons. By filling up the levels we can reproduce the *periodic table*. A discussion of this may be found in Supplement 14-B [www.wiley.com/college/gasiorowicz]. The configuration of the atoms from $Z = 1$ to $Z = 18$ are listed in the table.

Z	Element	Configuration	Term*	Ionization Potential eV
1	H	$(1s)$	$^2S_{1/2}$	13.6
2	He	$(1s)^2$	$^1S_0$	24.6
3	Li	$(\text{He})(2s)$	$^2S_{1/2}$	5.4
4	Be	$(\text{He})(2s)^2$	$^1S_0$	9.3
5	B	$(\text{He})(2s)^2(2p)$	$^2P_{1/2}$	8.3
6	C	$(\text{He})(2s)^2(2p)^2$	$^3P_0$	11.3
7	N	$(\text{He})(2s)^2(2p)^3$	$^4S_{3/2}$	14.5
8	O	$(\text{He})(2s)^2(2p)^4$	$^3P_2$	13.6
9	F	$(\text{He})(2s)^2(2p)^5$	$^2P_{3/2}$	17.4
10	Ne	$(\text{He})(2s)^2(2p)^6$	$^1S_0$	21.6
11	Na	$(\text{Ne})(3s)$	$^2S_{1/2}$	5.1
12	Mg	$(\text{Ne})(3s)^2$	$^1S_0$	7.6
13	Al	$(\text{Ne})(3s)^2(3p)$	$^2P_{1/2}$	6.0
14	Si	$(\text{Ne})(3s)^2(3p)^2$	$^3P_0$	8.1
15	P	$(\text{Ne})(3s)^2(3p)^3$	$^4S_{3/2}$	11.0
16	S	$(\text{Ne})(3s)^2(3p)^4$	$^3P_2$	10.4
17	Cl	$(\text{Ne})(3s)^2(3p)^5$	$^2P_{3/2}$	13.0
18	Ar	$(\text{Ne})(3s)^2(3p)^6$	$^1S_0$	15.8

*This describes the total spin, orbital, and J angular momentum of the ground state with the notation $^{2S+1}L_J$.

Whenever all the states for a given (n, l) are filled, we get a *closed shell* and the elements with $Z = 2, 4, 10, 12, 18, \dots$ are nonreactive, whereas elements with Z values ± 1 above these, which have one additional (*valence*) electron, or *hole* in a shell, interact strongly with elements that have holes and valence electrons. The shell structure is shown in Fig. 14-5. A somewhat more detailed description of the building up of the periodic table is reserved for Supplement 14-B.

A form of independent-electron approximation is used in simple molecular calculations. These are discussed in Supplement 14-C [www.wiley.com/college/gasiorowicz].

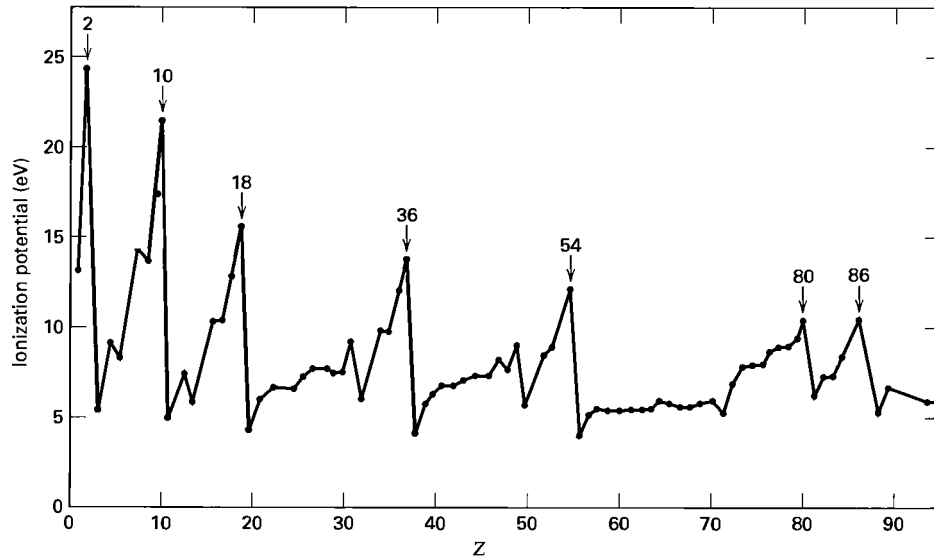


Figure 14-5 Ionization potential as a function of Z shows sharp peaks for closed shells from which electrons are hard to extract.

14-6 THE SIMPLEST MOLECULE, H_2^+

The simplest molecule consists of two hydrogen nuclei (protons of charge $Z = 1$) and a single electron. As in the case of helium, we are dealing with a three-body system, and a direct attack on the problem can only be carried out numerically. Here we also have a simple starting point: The basic approximation is the neglect of the motion of the protons. The fact that the proton is about 2000 times more massive than the electron justifies this. For our purposes, we therefore treat the protons as point charges that are infinitely heavy and that are located at the points $\pm \mathbf{R}/2$. On this basis, the energy eigenvalue equation for the electron is

$$\left(-\frac{\hbar^2}{2m_e} \nabla_r^2 - \frac{e^2}{4\pi\epsilon_0 |\mathbf{r} - \frac{\mathbf{R}}{2}|} - \frac{e^2}{4\pi\epsilon_0 |\mathbf{r} + \frac{\mathbf{R}}{2}|} + \frac{e^2}{4\pi\epsilon_0 R} \right) u(\mathbf{r}, \mathbf{R}) = E(R)u(\mathbf{r}, \mathbf{R}) \quad (14-50)$$

The first term is the kinetic energy operator for the electron; the next two terms represent the attraction of the electron to the two protons, and the last term represents the proton-proton repulsion. The qualitative features of the potential for fixed $\mathbf{R}$ are shown in Fig. 14-6. For very large R , the electron will be bound to one of the protons, and the energy of the system is -13.6 eV, the energy of a single hydrogen atom. When $R \rightarrow 0$, and we leave out the proton-proton repulsion, the electron will be bound to a $Z = 2$ nucleus, and the binding energy will be $-13.6Z^2$ eV. The electronic energy as a function of R interpolates smoothly between these values. When the energy of repulsion $e^2/4\pi\epsilon_0 R$ is added to this, the curve for $E(R)$ results. The curve dips below the level of the binding energy of an electron to a single proton, showing that the system of two protons and one electron is lower in energy than a proton and a separated hydrogen atom. Such a dip does not always exist, so that some atoms do not form molecules. To find $E(R)$ we must solve the Schrödinger equation (14-50). This can actually be done in elliptic coordinates, but we will get more insight from using the *variational principle* with trial wave functions that reflect some physical intuition about the system.

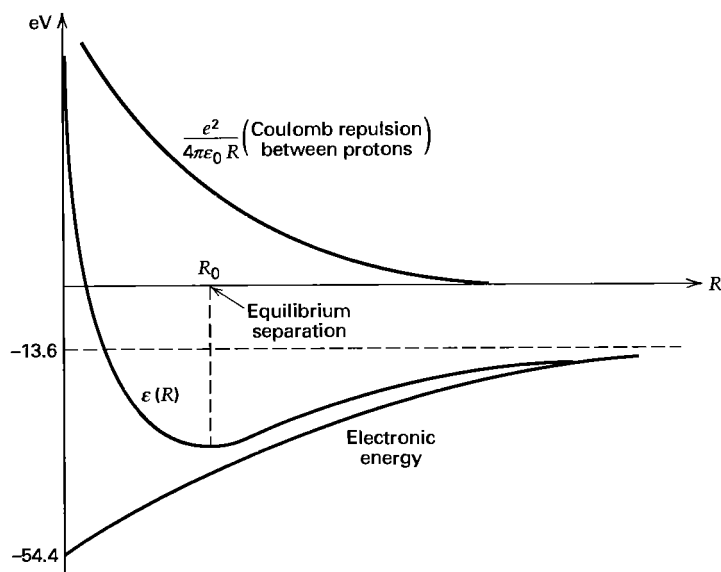


Figure 14-6 Contributions to “nuclear potential.” The Coulomb repulsion and the electronic energy combine to give a curve with a minimum at R_0 .

Molecular Orbitals

A reasonable trial wave function is a linear combination of $u_{100}(r_1, \mathbf{R})$ and $u_{100}(r_2, \mathbf{R})$, the ground-state wave functions of electrons bound to one or the other of the protons. Here $r_1 = |\mathbf{r} - \mathbf{R}/2|$ and $r_2 = |\mathbf{r} + \mathbf{R}/2|$. Since the two nuclei are identical, the Hamiltonian is symmetric under reflections in the plane bisecting the line joining them. This means that the Hamiltonian is invariant under reflections $\mathbf{r} \rightarrow -\mathbf{r}$. We may therefore choose linear combinations that are also eigenfunctions of parity, and therefore choose as our trial wave functions the so-called *molecular orbitals*

$$\begin{aligned}\psi_e(\mathbf{r}, \mathbf{R}) &= C_+(\psi_1 + \psi_2) \\ \psi_o(\mathbf{r}, \mathbf{R}) &= C_-(\psi_1 - \psi_2)\end{aligned}\quad (14-51)$$

where

$$\psi_{1,2} = \left(\frac{1}{\pi a_0^3}\right)^{1/2} e^{-|\mathbf{r} \mp \mathbf{R}/2|/a_0} \quad (14-52)$$

The normalization factors $C_{\pm}$ are given by

$$\frac{1}{C_{\pm}} = \langle \psi_1 \pm \psi_2 | \psi_1 \pm \psi_2 \rangle = 2 \pm 2 \int d^3r \psi_1(\mathbf{r}, \mathbf{R}) \psi_2(\mathbf{r}, \mathbf{R}) \quad (14-53)$$

The integral appearing above is called the *overlap integral*. The calculation of

$$\begin{aligned}S(R) &= \int d^3r \psi_1(\mathbf{r}, \mathbf{R}) \psi_2(\mathbf{r}, \mathbf{R}) \\ &= \frac{1}{\pi a_0^3} \int d^3r e^{-|\mathbf{r} - \mathbf{R}/2|/a_0} e^{-|\mathbf{r} + \mathbf{R}/2|/a_0} \\ &= \frac{1}{\pi a_0^3} \int d^3r' e^{-|\mathbf{r}' - \mathbf{R}|/a_0} e^{-r'/a_0}\end{aligned}\quad (14-54)$$

is straightforward, though tedious. The result is

$$S(R) = \left(1 + \frac{R}{a_0} + \frac{R^2}{3a_0^2}\right)e^{-R/a_0} \quad (14-55)$$

Notice that this quickly goes to zero when $R \gg a_0$.

The variational principle consists of calculating $\langle \psi_{e,o} | H | \psi_{e,o} \rangle$ as a function of R and minimizing this. We need to calculate

$$\begin{aligned} \langle H \rangle_{e,o} &= \frac{1}{2(1 \pm S(R))} \langle \psi_1 \pm \psi_2 | H | \psi_1 \pm \psi_2 \rangle \\ &= \frac{2\langle \psi_1 | H | \psi_1 \rangle \pm 2\langle \psi_1 | H | \psi_2 \rangle}{2(1 \pm S(R))} \end{aligned} \quad (14-56)$$

after use has been made of the symmetry under $\mathbf{R} \rightarrow -\mathbf{R}$. With

$$H = \frac{p_e^2}{2m_e} - \frac{e^2}{4\pi\epsilon_0|\mathbf{r} - \mathbf{R}/2|} - \frac{e^2}{4\pi\epsilon_0|\mathbf{r} + \mathbf{R}/2|} + \frac{e^2}{4\pi\epsilon_0 R} \quad (14-57)$$

we find that

$$\langle \psi_1 | H | \psi_1 \rangle = E_1 - \int d^3r \frac{e^2 |\psi_1(\mathbf{r}, \mathbf{R})|^2}{4\pi\epsilon_0|\mathbf{r} + \mathbf{R}/2|} + \frac{e^2}{4\pi\epsilon_0 R} \quad (14-58)$$

The first term is just the energy of a single hydrogen atom $E_1 = -13.6$ eV, the last term is the proton–proton repulsion, and the middle term is the electrostatic potential energy due to the electron charge distribution about one proton being attracted to the other proton. The integral can be evaluated, so that we finally get

$$\langle \psi_1 | H | \psi_1 \rangle = E_1 - \frac{e^2}{4\pi\epsilon_0 R} \left(1 + \frac{R}{a_0}\right) e^{-2R/a_0} \quad (14-59)$$

A more tedious calculation can be done to get

$$\langle \psi_1 | H | \psi_2 \rangle = \left(E_1 + \frac{e^2}{4\pi\epsilon_0 R}\right) S(R) - \frac{e^2}{4\pi\epsilon_0 a_0} \left(1 + \frac{R}{a_0}\right) e^{-R/a_0} \quad (14-60)$$

Here the last term is the exchange integral. When all of this is put together, the function to be minimized turns out to be

$$\langle H \rangle_{e,o} = \frac{1 - \frac{2(1+y)e^{-2y}}{y} \pm \left\{ \left(1 - \frac{2}{y}\right) \left(1 + y + \frac{y^2}{3}\right) e^{-y} - 2(1+y)e^{-y} \right\}}{1 \pm \left(1 + y - \frac{y^2}{3}\right) e^{-y}} \cdot E_1 \quad (14-61)$$

where $y = R/a_0$ and where we have made use of the fact that $e^2/4\pi\epsilon_0 a_0 = -2E_1$. Figure 14-7 is a plot of these energies as a function of R .

The exact solution, which according to the variational principle must lie below the curves obtained, differs little from the minimum. In our approximation, we see that the even solution yields binding, while the odd one does not. The difference between the even and the odd solutions is that in the former, the electron has a high probability of being located between the two protons, where the attractive contribution is maximized; for the

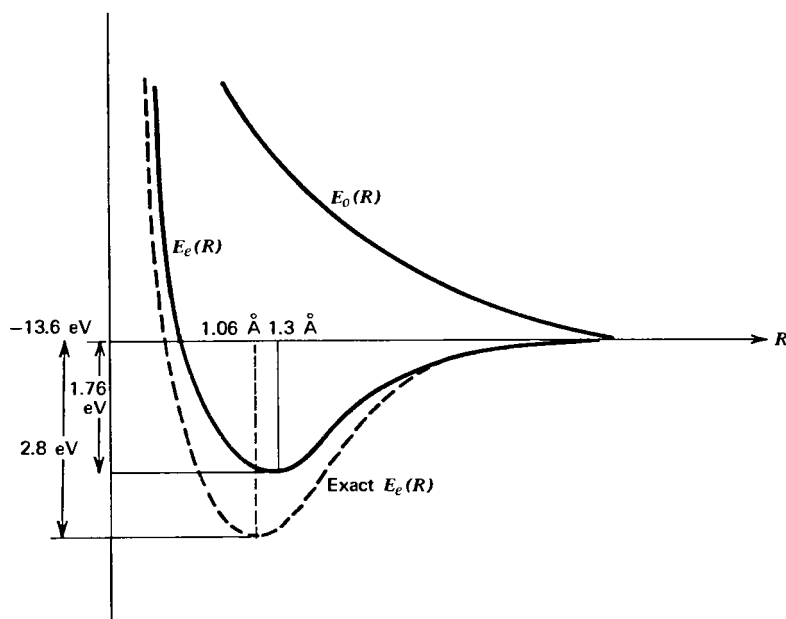


Figure 14-7 Results of variational calculation for H_2^+ .

odd solution, which has a node midway between the protons, the electron tends to be excluded from that region.

The experimental separation between the protons is 1.06 Å, and the binding energy is -2.8 eV. The calculations based on (14-61) lead to a separation of 1.3 Å and a binding energy of -1.76 eV. Thus our wave function is not as compact as it should be. The reason is that when R is small, the wave function should approach that of a He^+ ion, which (14-51) does not. One could improve the calculation by introducing an effective charge for the proton and minimizing $\langle H \rangle_e$ with respect to that parameter in addition to R , as in our illustration involving the helium atom. Since we are more interested in a qualitative understanding of the problem than in improving the variational calculation, we do not pursue this idea.

14-7 MOLECULAR SPECTRA

Our discussion of the H_2^+ molecule, as well as that of H_2 (in Supplement 14-C [www.wiley.com/college/gasiorowicz]) is based on the approximation that the nuclei are fixed in space, acting as a kind of scaffolding that upholds a cloud of charge, created by the rapidly moving electron (or electrons). Within that cloud of charge, the nuclei can still move. If we consider the function $E(R)$ as a potential in which the nuclei sit at the minimum, then any deviation from the equilibrium position corresponds to simple harmonic motion. If we write

$$E(R) \cong E(R_0) + \frac{1}{2} (R - R_0)^2 \left(\frac{d^2 E}{dR^2} \right)_{R=R_0} \quad (14-62)$$

we see a simple harmonic oscillator potential, centered at R_0 . The frequency of oscillation in that potential is given by ω , where

$$M\omega^2 = \left(\frac{d^2 E(R)}{dR^2} \right)_{R=R_0} \quad (14-63)$$

with M the reduced mass of the two nuclei. This potential gives rise to a vibrational spectrum, with energies

$$E_v = \hbar\omega\left(n_v + \frac{1}{2}\right) \quad n_v = 0, 1, 2, 3 \quad (14-64)$$

A rough estimate gives vibrational energies of the order of $\sqrt{m_e/M}$, the magnitude of the typical electronic energies (10 eV).⁷

There are even lower states of excitation. In the vibrational ground state, the molecule can still rotate. For a diatomic molecule, the rotation is about two axes perpendicular to the line joining the two nuclei. We may approximate the Hamiltonian by

$$H_{\text{rot}} = \frac{\mathbf{L}^2}{2I} \quad (14-65)$$

where I is the moment of inertia of the diatomic molecule. The spectrum is given by

$$E_{\text{rot}} = \frac{\hbar^2 L(L+1)}{2I} \quad (14-66)$$

The frequencies of radiation emitted in transitions between rotational levels allows us to determine I , and therefore the internuclear spacing. Since I is proportional to the nuclear mass, the rotational energies are of the order m_e/M of the electronic energies. This hierarchy of excitations shows up in curves of the specific heat of gases of molecules. The discussion of this is not within the scope of this book.

The Role of the Pauli Principle in Molecular Spectra

The lowest states of excitation of dipolar molecules are rotational ones. One treats the molecule as if it were a dumbbell, with two point masses separated by a distance R_{AB} . This system has two rotational degrees of freedom. If the internuclear separation defines the z -axis, then there are rotations about the x - and y -axes. There is no angular momentum about the z -axis, so that $L_z = 0$. Typically,

$$\begin{aligned} H_{\text{rot}} &= \frac{L_x^2 + L_y^2}{2I} = \frac{\mathbf{L}^2 - L_z^2}{2I} \rightarrow \\ E_{\text{rot}} &= \frac{L(L+1)\hbar^2}{2I} \end{aligned} \quad (14-67)$$

I is the moment of inertia of the molecule, and for homopolar molecules (identical nuclei) it is given by $I = \frac{1}{2}MR^2$, where M is the mass of each nucleus. Since the separation is of the order of a_0 , one sees that the energy differences are $(m/M) \times$ atomic energies. The point of interest for us is the fact that the Pauli principle has a big effect on molecular rotational spectra in homopolar molecules.

Consider, for example, the H_2 molecule for which the two nuclei are identical and each has spin $1/2$. Thus the total wave function must be antisymmetric under the interchange of the two nuclei. The two protons in this example may be in the antisymmetric spin singlet ($S = 0$) state, in which case the rotational state must be described by a symmetric function, so that the angular momentum is even. If the two protons are in the sym-

⁷For a simple discussion see J. Bernstein, P. M. Fishbane, and S. Gasiorowicz, *Modern Physics*, Prentice Hall, Englewood Cliffs, N.J., 2000.

metric spin triplet ($S = 1$) state, the angular momentum of rotation must be odd. In a gas, collisions among the H_2 molecules will randomize the distribution of spin states, and assuming that they have equal probability, the number of molecules in a given spin state will be proportional to the degeneracy ($2S + 1$). Thus there will be three times as many odd L molecules as there are even L H_2 molecules in the gas. This will manifest itself in the intensity of the spectral lines associated with the transitions between rotational levels. More generally, if each nucleus has spin I , then the spin states $2I, 2I - 2, 2I - 4, \dots$ and the spin states $2I - 1, 2I - 3, \dots$, will have opposite symmetry. If, for example, I is an integer, then the first series of spin states will be associated with even orbital angular momentum, since the nuclei are bosons in this case. Their total number is

$$\begin{aligned} \sum_{k=0}^I [2(2I - 2k) + 1] &= (4I + 1)(I + 1) - \frac{4I(I + 1)}{2} \\ &= (2I + 1)(I + 1) \end{aligned} \quad (14-68)$$

whereas the remaining

$$(2I + 1)^2 - (2I + 1)(I + 1) = (2I + 1)I$$

states will be associated with odd orbital angular momentum. Thus for integral I , the ratio of even L to odd L intensities for a given L is $(I + 1)/I$. For fermions that ratio is inverted.

From the study of rotational spectra one can identify the rotational levels and find their L values. The intensities then give a way of discriminating between even and odd spin. Historically, a study of the rotational spectrum of the N_2 molecule led to the conclusion that its spin was even. This could not be understood on the basis of a nuclear model in which the nitrogen nucleus consisted of fourteen protons and seven electrons; such a nucleus would have odd half-integral spin. The discovery of the neutron and the realization that the nitrogen nucleus consisted of seven protons and seven neutrons removed the difficulty.

PROBLEMS

1. Consider the helium atom in the approximation in which the electron–electron repulsion is neglected. What is the wave function for the ground state of *orthohelium* (spin 1)? What is its degeneracy in this approximation?
2. Write down and simplify the expression for the energy change due to the electron–electron repulsion in first-order perturbation theory. (Do not work out the final integrals!) Will this change the degeneracy you found in Problem 1? Estimate the order in which different levels occur.
3. Consider the lowest state of orthohelium as in Problems 1 and 2. Can you estimate the magnetic moment of that state?
4. Consider

$$“E” = \langle \Psi | H | \Psi \rangle$$

with an arbitrary trial wave function Ψ . Show that if Ψ differs from the correct ground-state wave function ψ_0 by terms of order ε , then “ E ” differs from the ground-state energy by terms of order ε^2 . [Note: Do not forget the normalization condition $\langle \Psi | \Psi \rangle = 1$]

5. Use the variational principle to estimate the ground-state energy of the three-dimensional harmonic oscillator, using the trial wave function

$$\Psi = N e^{-\alpha r}$$

6. Consider the Schrödinger equation for $l = 0$, with the potential given by

$$V(r) = V_0 f(r/r_0)$$

- (a) Show that it may be written in the form

$$\frac{d^2 u(x)}{dx^2} - \alpha^2 u(x) + \lambda f(x) u(x) = 0$$

where $u(0) = 0$ and that

$$\lambda = \frac{\int_0^\infty dx \left(\left(\frac{du}{dx} \right)^2 + \alpha^2 u^2(x) \right)}{\int_0^\infty dx f(x) u^2(x)}$$

- (b) Show that the correct eigenfunctions minimize λ .

7. Consider a finite-dimensional matrix H_{ij} . Show that the condition for minimizing

$$\langle \Psi | H | \Psi \rangle = \sum_{i,j=1}^n a_i^* H_{ij} a_j$$

subject to the condition

$$\langle \Psi | \Psi \rangle = \sum_{i=1}^n a_i^* a_i = 1$$

yields the eigenvalues of the matrix H . (*Hint:* Use the method of Lagrange multipliers.)

8. Use the variational principle to show that a one-dimensional attractive potential will always have a bound state. (*Hint:* Evaluate $\langle \Psi | H | \Psi \rangle$ with a convenient trial function—for example, $N e^{-\beta^2 x^2}$ —and show that this expectation value can always be made negative.)
9. Use the data of Fig. 14-4 to compute the location of the $(2s)(2p)$ level above the ground state of helium and compute the velocity of the electron emitted in autoionization, if the He^+ ion is in its lowest state at the end. What will it be if the He^+ ion is in its first excited state?
10. Consider a wave function $\psi(\alpha_1, \alpha_2, \dots, \alpha_n)$ for which only the dependence on some parameters is exhibited. The wave function is normalized

$$\langle \psi(\alpha_1, \alpha_2, \dots, \alpha_n) | \psi(\alpha_1, \alpha_2, \dots, \alpha_n) \rangle = 1$$

and the dependence on the parameters is so chosen that

$$\mathcal{E} = \langle \psi(\alpha_1, \dots) | H | \psi(\alpha_1, \dots) \rangle$$

is a minimum. Show that the parameters are determined by the set of equations

$$\left\langle \psi(\alpha_1, \dots) \left| H \left| \frac{\partial \psi}{\partial \alpha_i} \right. \right. \right\rangle - \mu \left\langle \psi(\alpha_1, \dots) \left| \frac{\partial \psi}{\partial \alpha_i} \right. \right\rangle = 0 \quad i = 1, 2, \dots, n$$

where μ is a Lagrange multiplier. Let H depend on a parameter λ (e.g., the nuclear charge or some distance—say, the internuclear distance in a molecule). Then the α_i will depend on that parameter. Prove that

$$\frac{d\mathcal{E}}{d\lambda} = \left\langle \psi(\alpha_1, \dots) \left| \frac{\partial H}{\partial \lambda} \right| \psi(\alpha_1, \dots) \right\rangle$$

This is known as the Feynman-Hellmann theorem and is very useful in molecular physics calculations.

11. Use the variational principle to estimate the ground-state energy for the anharmonic oscillator

$$H = \frac{p^2}{2m} + \lambda x^4$$

Compare your result with the exact result

$$E_0 = 1.060\lambda^{1/3} \left(\frac{\hbar^2}{2m} \right)^{2/3}$$

(Hint: Use a gaussian trial function.)

12. Use the exact result quoted in Problem 11 together with the Feynman-Hellmann theorem to calculate $\langle p^2 \rangle$ and $\langle x^4 \rangle$ for the ground state of the anharmonic oscillator.
13. According to the Ritz variational principle, the expectation of the Hamiltonian H in an arbitrary normalized state ψ obeys

$$\langle \psi | H | \psi \rangle > E_0$$

where E_0 is the lowest eigenvalue of H . Suppose H is an $N \times N$ hermitian matrix, whose elements are H_{ij} ($i, j = 1, 2, 3, \dots, N$) and E_0 is its lowest eigenvalue. Prove, using judicious choices of ψ , that E_0 is smaller than any one of the diagonal elements H_{ii} of the matrix H .

14. Consider two identical spin 1/2 particles in a harmonic oscillator potential, so that the Hamiltonian takes the form

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{1}{4} m \omega^2 (\mathbf{r}_1 - \mathbf{r}_2)^2$$

Assume that the two-particle system has zero center-of-mass momentum, and that the two particles are in $l = 0$ states.

- (a) Write down the ground-state wave function, including the spin state.
- (b) Write down the first excited states in the singlet and in the triplet spin states.
- (c) Assume that there is a short-range interaction between the particles, which in the $l = 0$ state can be approximated by $C [\delta(r)/r^2]$. Calculate the effect of this perturbation on the states obtained in (b).
15. The rotational motion of molecules has an effect on the equilibrium position of the nuclei. If R_0 is the separation for zero angular momentum, calculate the new separation by minimizing

$$E(R) = \frac{1}{2} M_{\text{red}} \omega^2 (R - R_0)^2 + \frac{\hbar^2 J(J+1)}{2M_{\text{red}} R^2}$$

assuming that the change from R_0 is small. How is the spectrum of the molecule affected? (Hint: The moment of inertia changes; calculate how.)

16. The $J = 0 \rightarrow J = 1$ absorption line in CO has a wavelength of 2.603 mm. Use this to calculate the moment of inertia of the CO molecule and the equilibrium separation.
17. Consider the H_2 molecule. The two nuclei (protons) have spin 1/2 and can therefore be in a total spin $S = 0$ or an $S = 1$ state.
- (a) What is the orbital angular momentum of the two-nucleon system in the lowest energy state for the two values of the total spin?
- (b) What is the wavelength of the radiation emitted in the transition from the lowest rotational excitation in the two cases?

Supplement 14-A

The Hartree Approximation

The energy eigenvalue problem for an atom with Z electrons has the form

$$\left[\sum_{i=1}^Z \left(\frac{\mathbf{p}_i^2}{2m_e} - \frac{Ze^2}{4\pi\epsilon_0 r_i} \right) + \sum_{i>j} \sum_j \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|} \right] \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) = E\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) \quad (14A-1)$$

and is a partial differential equation in $3Z$ dimensions. For light atoms it is possible to solve such an equation on a computer, but such solutions are only meaningful to the expert. We shall base our discussion of atomic structure on a different approach. As in the example of helium ($Z = 2$), it is both practical and enlightening to treat the problem as one involving Z independent electrons in a single potential, and to consider the electron–electron interaction later. Perturbation theory turned out to be adequate for $Z = 2$, but as the number of electrons increases, the shielding effects, not taken into account by first-order perturbation theory, become more and more important. The variational principle discussed at the end of Chapter 14 had the virtue of maintaining the single-particle picture, while at the same time yielding single-particle functions that incorporate the screening corrections.

To apply the variational principle, let us assume that the trial wave function is of the form

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_Z) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) \cdots \phi_Z(\mathbf{r}_Z) \quad (14A-2)$$

Each of the functions is normalized to unity. If we calculate the expectation value of H in this state, we obtain

$$\langle H \rangle = \sum_{i=1}^Z \int d^3r_i \phi_i^*(\mathbf{r}_i) \left(-\frac{\hbar^2}{2m_e} \nabla_i^2 - \frac{Ze^2}{4\pi\epsilon_0 r_i} \right) \phi_i(\mathbf{r}_i) + \frac{e^2}{4\pi\epsilon_0} \sum_{i>j} \sum_j \frac{|\phi_i(\mathbf{r}_i)|^2 |\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (14A-3)$$

The procedure of the variational principle is to pick the $\phi_i(\mathbf{r}_i)$ such that $\langle H \rangle$ is a minimum. If we were to choose the $\phi_i(\mathbf{r}_j)$ to be hydrogenlike wave functions, with a different Z_i for each electron (and with each electron in a different quantum state to satisfy the Pauli exclusion principle), we would get a set of equations analogous to (14-47) and (14-48). A more general approach is that due to Hartree. If the $\phi_i(\mathbf{r}_i)$ were the single-particle wave functions that minimized $\langle H \rangle$, then an alteration in these functions by an infinitesimal amount

$$\phi_i(\mathbf{r}_i) \rightarrow \phi_i(\mathbf{r}_i) + \lambda f_i(\mathbf{r}_i) \quad (14A-4)$$

should only change $\langle H \rangle$ by a term of order λ^2 . The alterations must be such that

$$\int d^3\mathbf{r}_i |\phi_i(\mathbf{r}_i) + \lambda f_i(\mathbf{r}_i)|^2 = 1$$

that is, to first order in λ ,

$$\int d^3\mathbf{r}_i [\phi_i^*(\mathbf{r}_i) f_i(\mathbf{r}_i) + \phi_i(\mathbf{r}_i) f_i^*(\mathbf{r}_i)] = 0 \quad (14A-5)$$

Let us compute the terms linear in λ that arise when (14A-4) is substituted into (14A-3). Term by term, we have

$$\begin{aligned} \sum_i \int d^3\mathbf{r}_i \left[\phi_i^*(\mathbf{r}_i) \left(-\frac{\hbar^2}{2m} \nabla_i^2 \right) \lambda f_i(\mathbf{r}_i) + \lambda f_i^*(\mathbf{r}_i) \left(-\frac{\hbar^2}{2m} \nabla_i^2 \right) \phi_i(\mathbf{r}_i) \right] \\ = \lambda \sum_i \int d^3\mathbf{r}_i \left\{ f_i(\mathbf{r}_i) \left[-\frac{\hbar^2}{2m} \nabla_i^2 \phi_i^*(\mathbf{r}_i) \right] + f_i^*(\mathbf{r}_i) \left[-\frac{\hbar^2}{2m} \nabla_i^2 \phi_i(\mathbf{r}_i) \right] \right\} \end{aligned} \quad (14A-6)$$

To obtain this we have integrated by parts two times, and used the fact that $f_i(\mathbf{r}_i)$ must vanish at infinity in order to be an acceptable variation of a square integrable function. Next we have

$$-\lambda \sum_i \int d^3\mathbf{r}_i \left[f_i^*(\mathbf{r}_i) \frac{Ze^2}{4\pi\epsilon_0 r_i} \phi_i(\mathbf{r}_i) + \phi_i^*(\mathbf{r}_i) \frac{Ze^2}{4\pi\epsilon_0 r_i} f_i(\mathbf{r}_i) \right] \quad (14A-7)$$

and finally

$$\frac{\lambda e^2}{4\pi\epsilon_0} \sum_{i>j} \sum_j \int d^3\mathbf{r}_i \int d^3\mathbf{r}_j \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} [f_i^*(\mathbf{r}_i) \phi_i(\mathbf{r}_i) + f_i(\mathbf{r}_i) \phi_i^*(\mathbf{r}_i)] |\phi_j(\mathbf{r}_j)|^2 + (i \leftrightarrow j) \quad (14A-8)$$

We cannot just set the sum of these three terms equal to zero because the $f_i(\mathbf{r}_i)$ are constrained by (14A-5). The proper way to account for the constraint is by the use of Lagrange multipliers; that is, we multiply each of the constraining relations (14A-5) by a constant (the “multiplier”) and add the sum to our three terms. The total can then be set equal to zero, since the constraints on the $f_i(\mathbf{r}_i)$ are now taken care of. With a certain amount of notational foresight we label the multipliers $-\epsilon_i$, and thus get

$$\begin{aligned} \sum_i \int d^3\mathbf{r}_i \left\{ f_i^*(\mathbf{r}_i) \left[-\frac{\hbar^2}{2m} \nabla_i^2 \phi_i(\mathbf{r}_i) - \frac{Ze^2}{4\pi\epsilon_0 r_i} \phi_i(\mathbf{r}_i) \right] \right\} \\ + \frac{e^2}{4\pi\epsilon_0} \sum_{i>j} \sum_j \int \int d^3\mathbf{r}_i d^3\mathbf{r}_j f_i^*(\mathbf{r}_i) \frac{|\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \phi_i(\mathbf{r}_i) \\ - \epsilon_i \int d^3\mathbf{r}_i f_i^*(\mathbf{r}_i) \phi_i(\mathbf{r}_i) + \text{complex conjugate terms} = 0 \end{aligned} \quad (14A-9)$$

In deriving the second line, first we converted the double sum $\sum_{i>j} \sum_j$ into $(1/2) \sum_{i \neq j} \sum_j$, which is unrestricted except for the requirement that $i \neq j$, and then used the fact that the integrand in (14A-8) is symmetric in i and j . Now $f_i(\mathbf{r}_i)$ is completely unrestricted, so that we may treat $f_i(\mathbf{r}_i)$ and $f_i^*(\mathbf{r}_i)$ as completely independent (each one has a real and an imaginary part). Furthermore, other than being square integrable, they are completely arbitrary, so that for (14A-5) to hold, the coefficients of $f_i(\mathbf{r}_i)$ and $f_i^*(\mathbf{r}_i)$ must separately vanish *at each point* $\mathbf{r}_i$, since we are allowed to make local variations in the functions $f_i(\mathbf{r}_i)$ and $f_i^*(\mathbf{r}_i)$. We are thus led to the condition that

$$\left[-\frac{\hbar^2}{2m_e} \nabla_i^2 - \frac{Ze^2}{4\pi\epsilon_0 r_i} + \frac{e^2}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3\mathbf{r}_j \frac{|\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \phi_i(\mathbf{r}_i) = \epsilon_i \phi_i(\mathbf{r}_i) \quad (14A-10)$$

and the complex conjugate relation.

This equation has a straightforward interpretation: It is an energy eigenvalue equation for electron “ i ” located at $\mathbf{r}_i$, moving in a potential

$$V_i(\mathbf{r}_i) = -\frac{Ze^2}{4\pi\epsilon_0 r_i} + \frac{e^2}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3r_j \frac{|\phi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (14A-11)$$

that consists of an attractive Coulomb potential due to a nucleus of charge Z , and a repulsive contribution due to the charge density of all the other electrons. We do not, of course, know the charge densities

$$\rho_j(\mathbf{r}_j) = -e|\phi_j(\mathbf{r}_j)|^2$$

of all the other electrons, so that we must search for a *self-consistent* set of $\phi_i(\mathbf{r}_i)$, in the sense that their insertion in the potential leads to eigenfunctions that reproduce themselves. The equation (14A-10) is a rather complicated integral equation, but it is at least an equation in three dimensions (we can replace the variable $\mathbf{r}_i$ by $\mathbf{r}$), and that makes numerical work much easier. An even greater simplification occurs when $V_i(\mathbf{r})$ is replaced by its angular average

$$V_i(r) = \int \frac{d\Omega}{4\pi} V_i(\mathbf{r}) \quad (14A-12)$$

for then the self-consistent potential becomes central, and the self-consistent solutions can be decomposed into angular and radial functions; that is, they will be functions that can be labeled by n_i, l_i, m_i, σ_i , with the last label referring to the spin state ($s_{iz} = \pm 1/2$).

The trial wave function (14A-2) does not take into account the exclusion principle. The latter plays an important role, since if all the electrons could be in the same quantum state, the energy would be minimum with all the electrons in the $n = 1, l = 0$ “orbital.” Atoms do not have such a simple structure. To take the exclusion principle into account, we add to the *Ansatz* represented by (14A-2) the rule: *Every electron must be in a different state*, if the spin states are included in the labeling. A more sophisticated way of doing this automatically is to replace (14A-2) by a trial wave function that is a *Slater determinant* [cf. Eq. (13-39)]. The resulting equations differ from (14A-10) by the addition of an exchange term. The new Hartree-Fock equations have eigenvalues that turn out to differ by 10–20 percent from those obtained using Hartree equations supplemented by the condition arising from the exclusion principle. It is a little easier to talk about the physics of atomic structure in terms of the Hartree picture, so we will not discuss the Hartree-Fock equations.

The potential (14A-12) no longer has the $1/r$ form, and thus the degeneracy of all states with a given n and $l \leq n - 1$ is no longer present. We may expect, however, that for low Z at least, the splitting for different l values for a given n will be smaller than the splitting between different n -values, so that electrons placed in the orbitals $1s, 2s, 2p, 3s, 3p, 3d, 4s, 4p, 4d, 4f, \dots$ will be successively less strongly bound. Screening effects will accentuate this: Whereas s orbitals do overlap the small r region significantly, and thus feel the full nuclear attraction, the p -, d -, $\dots$ orbitals are forced out by the centrifugal barrier, and feel less than the full attraction. This effect is so strong that the energy of the $3d$ electrons is very close to that of the $4s$ electrons, so that the anticipated ordering is sometimes disturbed. The same is true for the $4d$ and $5s$ electrons, the $4f$ and $6s$ electrons, and so on. The dominance of the l -dependence over the n -dependence becomes more important as we go to larger Z -values, as we shall see in our discussion of the *periodic table*.

W-54 Supplement 14-A The Hartree Approximation

The number of electrons that can be placed in orbitals with a given (n, l) is $2(2l + 1)$, since there are two spin states for given m -value. When all these $2(2l + 1)$ states are filled, we speak of the *closing of a shell*. The charge density for a closed shell has the form

$$-e \sum_{m=-l}^l |R_{nl}(r)|^2 |Y_{lm}(\theta, \phi)|^2 \quad (14A-13)$$

and this is spherically symmetric because of the property of spherical harmonics that

$$\sum_{m=-l}^l |Y_{lm}(\theta, \phi)|^2 = \frac{2l + 1}{4\pi} \quad (14A-14)$$

Supplement 14-B

The Building-Up Principle

In this section we discuss the building up of atoms by the addition of more and more electrons to the appropriate nucleus, whose only role, to good approximation, is to provide the positive charge Ze .

Hydrogen ($Z = 1$) There is only one electron, and the ground-state configuration is $(1s)$. The ionization energy is 13.6 eV, and the amount of energy needed to excite the first state above the ground state is 10.2 eV. The radius of the atom is 0.5 Å, and its spectroscopic description is $^2S_{1/2}$.

Helium ($Z = 2$) The lowest two-electron state, as we saw in Chapter 14, is one in which both electrons are in the $(1s)$ orbital. We denote this configuration by $(1s)^2$. In spectroscopic notation, the ground state is an $l = 0$ spin singlet state, $^1S_0$ because the exchange effect favors it. The total binding energy is 79 eV. After one electron is removed, the remaining electron is in a $(1s)$ orbit about a $Z = 2$ nucleus. Thus its binding energy is $13.6Z^2$ eV = 54.4 eV, and the energy required to remove the first electron, the *ionization energy*, is $79.0 - 54.4 = 24.6$ eV. A rough estimate of the energy of the first excited state, with configuration $(1s)(2s)$, is $-13.6Z^2 - 13.6(Z - 1)^2/n^2 \approx -58$ eV for $Z = 2$ and $n = 2$. This expression takes into account shielding in the second term. Thus the excitation energy is $79 \text{ eV} - 58 \text{ eV} \approx 21 \text{ eV}$.¹ In any reaction with another substance, about 20 eV is required for a rearrangement of the electrons, and thus helium is chemically very inactive. This property is shared by all atoms whose electrons form closed shells, but the energy required is particularly large for helium.

Lithium ($Z = 3$) The exclusion principle forbids a $(1s)^3$ configuration, and the lowest energy electron configuration is $(1s)^2(2s)$. We are thus adding an electron to a closed shell, and since the shell is in a $^1S_0$ state, the spectroscopic description of the ground state is $^2S_{1/2}$, just as for hydrogen. If the screening were perfect, we should expect a binding energy of -3.4 eV (since $n = 2$). The screening is not perfect, especially since the outer *valence electron* being in an s -state, its wave function has a reasonable overlap with nucleus at $r = 0$. We can estimate the effective Z from the measured ionization energy of 5.4 eV, and it is $Z^* = 1.3$. It takes very little energy to excite the lithium atom. The six $(2p)$ electronic states lie just a little above the $(2s)$ state, and these $(2p)$ states, when occupied, make the atom chemically active (see our more extended discussion of carbon). Lithium, like other elements that have one electron outside a closed shell, is a very active element.

¹This is a crude estimate that ignores the electron–electron repulsion and exchange effects. The difference between the 21 eV and the 24.6 eV is the 4–5 eV that will be released when the excited atom decays to its ground state. (See Fig. 14-2b.)

Beryllium ($Z = 4$) The natural place for the fourth electron to go is into the second space in the (2s) orbital, so that the configuration is $(1s)^2(2s)^2$. We again have a closed shell and the spectroscopic description is $^1S_0$. As far as the energy is concerned, the situation is very much like that of helium. If the screening were perfect, we might expect a binding energy like that of helium, since the inner electrons reduce the effective Z to something like $Z = 2$. Since $n = 2$, we would expect an ionization energy of $24.6/n^2 = 6.2$ eV. The shielding situation is somewhat like that for lithium, and if we make a guess that, as in lithium, the binding energy is increased by about 50 percent, we get approximately 9 eV. The experimental value is 9.3 eV. Although the shell is closed, the excitation of one of the electrons to a (2p) orbital does not cost much energy. Thus, in the presence of another element a rearrangement of electrons may yield enough energy to break up the closed shell. We therefore expect beryllium not to be as inert as helium. It is generally true that atoms in which the outer electrons have their spins “paired up” into singlet states are less reactive.

Boron ($Z = 5$) After the closing of the shell, the fifth electron can either go into a (3s) orbital or into a (2p) orbital. The latter is lower in energy, and it is the (2p) shell that begins to fill up, starting with boron. The configuration is $(1s)^2(2s)^2(2p)$, and the spectroscopic description of the state is $^2P_{1/2}$. This deserves comment: If we add spin 1/2 to an $l = 1$ orbital state, we may have $J = 3/2$ or $1/2$. These states are split by a spin-orbit interaction

$$\frac{1}{2m_e^2c^2} \mathbf{L} \cdot \mathbf{S} \frac{1}{r} \frac{dV(r)}{dr} = \frac{\hbar^2}{4m_e^2c^2} [J(J+1) - L(L+1) - 3/4] \frac{1}{r} \frac{dV(r)}{dr} \quad (14B-1)$$

and the form of this leads to the higher J value having a higher energy, because the expectation value of $(1/r)(dV/dr)$ is still positive, even though it is no longer equal to the value given in (12-16). This conclusion depends on the degree to which the shell is filled, as specifically given by *Hund's rules*. These will be discussed below. The ionization energy is 8.3 eV. This meets the expectation that the value should be somewhat lower than that for beryllium, since the 2p state energy is somewhat higher than that of the 2s orbital.

Carbon ($Z = 6$) The configuration for carbon is $(1s)^2(2s)^2(2p)^2$. The second electron could be in the same p -state as the first electron, with the two of them making an up-down spin pair. It is, however, advantageous for the second to stay out of the way of the first electron, thus lowering the repulsion between the electrons. It can do so because the possible $l = 1$ states Y_{11} , Y_{10} , Y_{1-1} allow for the linear combinations $\sin \theta \cos \phi$, $\sin \theta \sin \phi$, and $\cos \theta$, which are aligned along the x -, y -, and z -axes, respectively (Fig. 14B-1). When two electrons go into orthogonally aligned arms, the overlap is minimized and the repulsion is reduced. The electrons are in different spatial states, so that their spins do not have to be antiparallel. One might expect carbon to be divalent. This is not so, because of the subtleties that arise from close-lying energy levels. It costs very little energy to promote one of the (2s) electrons into the third unoccupied $l = 1$ state. The configuration $(1s)^2(2s)(2p)^3$ has four “unpaired” electrons, and the gain in energy from the formation of four bonds with other atoms more than makes up for the energy needed to promote the (2s) electron. The reduction in the repulsion leads to a somewhat larger ionization energy than that for boron, 11.3 eV. The spectroscopic description of the ground state is $^3P_0$. We can have a total spin of 0 or 1 for the two 2p electrons, and, since we are adding two $l = 1$ states, the total orbital angular momentum can be 0, 1, or 2. Of the various states, $^1S_0$, $^3P_{2,1,0}$, and $^1D_2$, the state of higher spin has the lower energy (cf. our discussion of helium) and by another of *Hund's rules*, the $^3P_0$ state has the lowest energy.

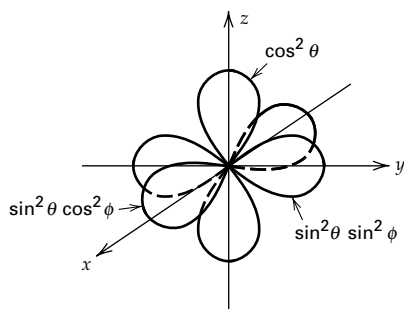


Figure 14B-1 Orthogonal distributions of electronic clouds in the $l = 1$ state.

Nitrogen ($Z = 7$) Here the configuration is $(1s)^2(2s)^2(2p)^3$, sometimes described as $(2p)^3$ for brevity (the closed shells and subshells are omitted). The three electrons can all be in nonoverlapping p -states, and thus we expect the increase in ionization energy to be the same as the increase from boron to carbon. This is in agreement with the measured value of 14.5 eV.

Oxygen ($Z = 8$) The configuration may be abbreviated to $(2p)^4$, and the shell is more than half full. Since there are four electrons, it appears as if the determination of the ground-state spectroscopic state would be very difficult. We can, however, look at the shell in another way. We know that when two more electrons are added to make a $(2p)^6$ configuration, then the shell is filled, and the total state has $L = S = 0$. We can thus think of oxygen as having a closed $2p$ shell with two *holes* in it. These holes are just like anti-electrons, and we can look at two-hole configurations. These will be the same as two-electron configurations, since the holes also have spin $1/2$. Thus, as with carbon, the possible states consistent with the antisymmetry of the two-fermion (two-hole) wave function are ^1S , ^3P , ^1D , and the four electrons must be in the same states, since they, together with the two-hole system, give $S = 0$, $L = 0$. The highest spin is $S = 1$, and thus we must have a ^3P state. Hund's rule, which will be discussed in the next section, yields the $^3P_2$ state. When the fourth electron is added to the nitrogen configuration, it must go into an orbital with an m -value already occupied. Thus two of the electron wave functions overlap, and this raises the energy because of the repulsion. It is therefore not surprising that the ionization energy drops to the value of 13.6 eV.

Fluorine ($Z = 9$) Here the configuration is $(2p)^5$. The monotonic increase in the ionization energy resumes, with the experimental value of 17.4 eV. Fluorine is chemically very active, because it can "accept" an electron to form a closed shell $(2p)^6$, which is very stable. Since the addition of a single electron with $s = 1/2$ and $l = 1$ yields a $^1S_0$ state, the shell with the hole in it must have $s = 1/2$ and $l = 1$. It is therefore a ^2P state, and by Hund's rule, as we shall see, the state is $^2P_{3/2}$.

Neon ($Z = 10$) With $Z = 10$ the $(2p)$ shell is closed, and all electrons are paired off. The ionization energy is 21.6 eV, continuing the monotonic trend. Here, as in helium, the first available state that an electron can be excited into has a higher n value, and thus it takes quite a lot of energy to perturb the atom. Neon, like helium, is an inert gas.

At this point, the addition of another electron requires putting it in an orbit with a higher n value ($n = 3$), and thus neon marks the end of a *period* in the periodic table, as did helium. In neon, as in helium, the first available state into which an electron can be excited has a higher n -value, so that it takes quite a lot of energy to perturb the atom. Neon shares with helium the property of being an *inert gas*.

The next period again has eight elements in it. First the (3s) shell is filled, with sodium ($Z = 11$) and magnesium ($Z = 12$), and then the 3p shell, which includes, in order, aluminum ($Z = 13$), silicon ($Z = 14$), phosphorus ($Z = 15$), sulphur ($Z = 16$), chlorine ($Z = 17$) and, closing the shell, argon ($Z = 18$). These elements are chemically very much like the series: lithium, . . . , neon, and the spectroscopic descriptions of the ground states are the same. The only difference is that, since $n = 3$, the ionization energies are somewhat smaller, as can be seen from the periodic table at the end of this supplement.

It might appear a little strange that the period ends with argon, since the (3d) shell, accommodating ten elements, remains to be filled. The fact is that the self-consistent potential is not of the $1/r$ form, and the intrashell splitting here is sufficiently large that the (4s) state lies lower than the (3d) state, though not by much. Hence a competition develops, and in the next period we have (4s), (4s)², (4s)²(3d), (4s)²(3d)², (4s)²(3d)³, (4s)(3d)⁵, (4s)²(3d)⁵, (4s)²(3d)⁶, (4s)²(3d)⁷, (4s)²(3d)⁸, (4s)(3d)¹⁰, (4s)²(3d)¹⁰ and then the 4p shell gets filled until the period ends with krypton ($Z = 36$). The chemical properties of elements at the beginning and end of this period are similar to those of elements at the beginning and end of other periods. Thus potassium, with the single (4s) electron, is an alkali metal, like sodium with its single (3s) electron outside a closed shell. Bromine, with the configuration (4s)²(3d)¹⁰(4p)⁵, has a single hole in a p-shell and thus is chemically like chlorine and fluorine. The series of elements in which the (3d) states are being filled all have rather similar chemical properties. The reason for this again has to do with the details of the self-consistent potential. It turns out that the radii of these orbits² are somewhat smaller than those of the (4s) electrons, so that when the (4s)² shell is filled, these electrons tend to shield the (3d) electrons, no matter how many there are, from outside influences. The same effect occurs when the (4f) shell is being filled, just after the (6s) shell has been filled. The elements here are called the *rare earths*.

Spectroscopic Description of Ground States

In our discussion of the light atoms, we often gave the spectroscopic description of the ground states—for example, ³P₂ for oxygen, ²P_{3/2} for fluorine, and so on. The knowledge of S , L , and J for the ground states is important, because selection rules allow us to determine these quantities for the excited states of atoms. We referred to *Hund's rules* in the determination of these, and these rules are the subject of this section.

What determines the ground-state quantum numbers is an interplay of spin-orbit coupling and the exchange effect discussed in connection with helium in Chapter 14. For the lighter atoms ($Z < 40$), for which the motion of the electrons is nonrelativistic, the electron–electron repulsion effects are more important than the spin-orbit coupling. This means that it is a fairly good approximation to view L and S as separately good quantum numbers: We add up all the spins to form an S and all the orbital angular momenta of the electrons to form an L , and these are then coupled to obtain a total J . For heavier atoms it is a better approximation to first couple the spin and orbital angular momentum to form a total angular momentum for that electron, and then to couple all of the J 's together. The former case is described as Russell-Saunders coupling, the latter as j - j coupling. For Russell-Saunders coupling, F. Hund summarized the results of

²It is understood that this is just a way of talking about the peaking tendencies of the charge distribution.

various calculations by a set of rules that give the overall quantum numbers of the lowest states. The rules are:

1. The state with largest S lies lowest.
2. For a given value of S , the state with maximum L lies lowest.
3. For a given L and S , if the incomplete shell is not more than half-filled, the lowest state has the minimum value of $J = |L - S|$; if the shell is more than half-filled, the state of lowest energy has $J = L + S$.

In applying these rules we must be careful not to violate the Pauli principle.

The first of these rules is easy to understand: The largest S state is symmetric in all the spins (since it contains the state $S_z = S_{\max}$ for which all the spins are parallel), and thus the spatial wave function is antisymmetric, which minimizes the electron overlap, and thus the expectation value of the repulsive potentials.

The second rule emerges qualitatively from the fact that the higher the L -value, the more lobes the wave function has, as shown in Fig. 8-4. This allows the electrons to stay away from one another, and reduce the effect of the Coulomb repulsion.

The third rule follows from the form of the spin-orbit coupling. Since the expectation value of $[(1/r)(dV/dr)]$ is positive, the perturbation due to the spin-orbit coupling splits the degenerate J states (for a given L and S) and it is clear from (14B-1) that the lowest value of J will give the lowest lying state. Once we get to the point of having a shell that is more than half filled, it is clearer to look at the atom as consisting of a filled shell with a number of holes, as we discussed in our description of oxygen. These holes act as if they had positive charge, and for the spin-orbit interaction of the holes, the sign of $[(1/r)(dV/dr)]$ is reversed. Thus the multiplet is *inverted* and it is the largest value of J that gives the lowest lying state.

Let us illustrate the application of the Hund rules to some atoms, and the need to keep track of the Pauli principle. We shall consider the quantum numbers of carbon $(2p)^2$, oxygen $(2p)^4$, and manganese $(3d)^5$. In the first two cases we have p -states, so that we can draw a set of "shelves" corresponding to $L_z = 1, 0, -1$. The electrons are placed, as far as is possible, on different shelves, to minimize the repulsion. For carbon we place them in $L_z = 1$ and $L_z = 0$ states. By Hund's first rule, the spins will be parallel (strictly speaking they will be in a triplet state) (Fig. 14B-2). Thus we have $S_z = 1$ for the largest possible value, and we get a triplet state. The largest possible value of L_z gives the L value, which is 1. The third rule thus gives $J = |L - S| = 0$, and we have a $^3P_0$ state. For the $(2p)^4$ case, we fill all three shelves with one electron each, and then put the last electron in the $L_z = 1$ state, for example. The Pauli principle demands that the two electrons in the $L_z = 1$ state form a singlet. Thus only the other two electrons are relevant, and since $S_z = 1$, $S = 1$. The maximum value of $L_z = [2 + 0 + (-1)] = 1$, so that $L = 1$. Now, however, we have more than a half-filled shell, so that $J = L + S = 2$, and oxygen has a $^3P_2$ ground state.

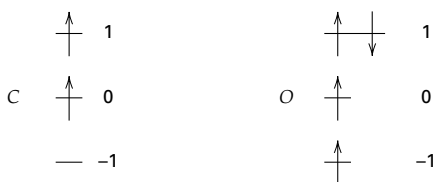


Figure 14B-2 Application of Hund's rules to six electron system.

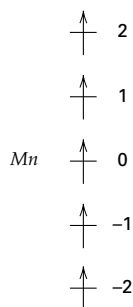


Figure 14B-3 Application of Hund's rules to atom with five valence electrons in d -state.

For manganese the shelves have $L_z = 2, 1, 0, -1, -2$, as shown in Fig. 14B-3. There are five electrons, and thus each of the spaces is filled by one. With the spins parallel we get $S_z = 5/2$, which implies that $S = 5/2$. The total value of $L_z = 0$, and thus we have an S -state. This means that the ground state is an ${}^6S_{5/2}$ state.

Limitations of space prevent us from a more detailed discussion of the periodic table. A few additional comments are, however, in order.

- (a) There is nothing in atomic structure that limits the number of elements. The reason that atoms with $Z \gtrsim 100$ do not occur naturally is that heavy *nuclei* undergo spontaneous fission. If new, superheavy (metastable) nuclei are ever discovered, there will presumably exist corresponding atoms, and it is expected that their structure will conform to the prediction of the building-up approach outlined in this supplement.
- (b) The ionization energies all lie in the vicinity of 5–15 eV. The reason is that in spite of the increasing number of electrons, the outermost electrons “see” a charge that lies in the range $Z = 1$ –2. In addition, because of the departures from a point charge distribution, the dependence of the energy is no longer of the $1/n^2$ form. Consequently, the wave functions of the outermost electrons do not extend much further than that of the electron in the hydrogen atom. Atoms are more or less the same size!
- (c) We went to a great deal of trouble to specify the S , L , and J quantum numbers of the ground states of the various elements. The reason for doing this is that in spectroscopy, the quantum numbers are of particular interest because of the selection rules

$$\begin{aligned}\Delta S &= 0 \\ \Delta L &= \pm 1 \\ \Delta J &= 0, \pm 1 \quad (\text{no } 0 \rightarrow 0)\end{aligned}\tag{14B-2}$$

that will be derived later, and that may then be used to determine the quantum numbers of the excited states. The spectroscopy of atoms, once we get beyond hydrogen and helium, is very complicated. Consider, as a relatively simple example, the first few states of carbon, which are formed from different configurations of the two electrons that lie outside the closed shell in the $(2p)^2$ orbitals. As already pointed out, the possible states are ${}^1S_0$, ${}^3P_{2,1,0}$, and ${}^1D_2$. The ${}^3P_0$ state lies lowest, but the other states are still there. The first excited states may be described by the orbitals $(2p)(3s)$. Here $S = 0$ or 1, but $L = 1$ only. Since the n -values are different, the exclusion principle does not restrict the states in any way, and all of the states ${}^1P_1$, ${}^3P_{2,1,0}$ are possible, while the excited states that

arise from the orbitals $(2p)(3p)$ can have $S = 0, 1$ and $L = 2, 1, 0$, leading to all the states $^1D_2$, $^1P_1$, $^1S_0$, $^3D_{3,2,1}$, $^3P_{2,1,0}$, and $^3S_1$. Even with the restrictions provided by the selection rules, there are numerous transitions. Needless to say, the ordering of these levels represents a delicate balance between various competing effects, and the prediction of the more complex spectra is very difficult. That task is not really of interest to us, since the main point that we want to make is that quantum mechanics provides a qualitative, and quantitative, detailed explanation of the chemical properties of atoms and of their spectra, without assuming an interaction other than the electromagnetic interaction between charged particles. We shall have occasion to return to the topic of spectra.

Periodic Table

Z	Element	Configuration	Term ¹	Ionization Potential eV	Radius ² in Å
1	H	(1s)	$^2S_{1/2}$	13.6	0.53
2	He	(1s) ²	$^1S_0$	24.6	0.29
3	Li	(He)(2s)	$^2S_{1/2}$	5.4	1.59
4	Be	(He)(2s) ²	$^1S_0$	9.3	1.04
5	B	(He)(2s) ² (2p)	$^2P_{1/2}$	8.3	0.78
6	C	(He)(2s) ² (2p) ²	$^3P_0$	11.3	0.62
7	N	(He)(2s) ² (2p) ³	$^4S_{3/2}$	14.5	0.52
8	O	(He)(2s) ² (2p) ⁴	$^3P_2$	13.6	0.45
9	F	(He)(2s) ² (2p) ⁵	$^2P_{3/2}$	17.4	0.40
10	Ne	(He)(2s) ² (2p) ⁶	$^1S_0$	21.6	0.35
11	Na	(Ne)(3s)	$^2S_{1/2}$	5.1	1.71
12	Mg	(Ne)(3s) ²	$^1S_0$	7.6	1.28
13	Al	(Ne)(3s) ² (3p)	$^2P_{1/2}$	6.0	1.31
14	Si	(Ne)(3s) ² (3p) ²	$^3P_0$	8.1	1.07
15	P	(Ne)(3s) ² (3p) ³	$^4S_{3/2}$	11.0	0.92
16	S	(Ne)(3s) ² (3p) ⁴	$^3P_2$	10.4	0.81
17	Cl	(Ne)(3s) ² (3p) ⁵	$^2P_{3/2}$	13.0	0.73
18	Ar	(Ne)(3s) ² (3p) ⁶	$^1S_0$	15.8	0.66
19	K	(Ar)(4s)	$^2S_{1/2}$	4.3	2.16
20	Ca	(Ar)(4s) ²	$^1S_0$	6.1	1.69
21	Sc	(Ar)(4s) ² (3d)	$^2D_{3/2}$	6.5	1.57
22	Ti	(Ar)(4s) ² (3d) ²	$^3F_2$	6.8	1.48
23	V	(Ar)(4s) ² (3d) ³	$^4F_{3/2}$	6.7	1.40
24	Cr	(Ar)(4s)(3d) ⁵	$^7S_3$	6.7	1.45
25	Mn	(Ar)(4s) ² (3d) ⁵	$^6S_{5/2}$	7.4	1.28
26	Fe	(Ar)(4s) ² (3d) ⁶	$^5D_4$	7.9	1.23
27	Co	(Ar)(4s) ² (3d) ⁷	$^4F_{9/2}$	7.8	1.18
28	Ni	(Ar)(4s) ² (3d) ⁸	$^3F_4$	7.6	1.14
29	Cu	(Ar)(4s)(3d) ¹⁰	$^2S_{1/2}$	7.7	1.19
30	Zn	(Ar)(4s) ² (3d) ¹⁰	$^1S_0$	9.4	1.07
31	Ga	(Ar)(4s) ² (3d) ¹⁰ (4p)	$^2P_{1/2}$	6.0	1.25

(Continued)

Periodic Table (Continued)

Z	Element	Configuration	Term ¹	Ionization Potential eV	Radius ² in Å
32	Ge	(Ar)(4s) ² (3d) ¹⁰ (4p) ²	³ P ₀	8.1	1.09
33	As	(Ar)(4s) ² (3d) ¹⁰ (4p) ³	⁴ S _{3/2}	10.0	1.00
34	Se	(Ar)(4s) ² (3d) ¹⁰ (4p) ⁴	³ P ₂	9.8	0.92
35	Br	(Ar)(4s) ² (3d) ¹⁰ (4p) ⁵	² P _{3/2}	11.8	0.85
36	Kr	(Ar)(4s) ² (3d) ¹⁰ (4p) ⁶	¹ S ₀	14.0	0.80
37	Rb	(Kr)(5s)	² S _{1/2}	4.2	2.29
38	Sr	(Kr)(5s) ²	¹ S ₀	5.7	1.84
39	Y	(Kr)(5s) ² (4d)	² D _{3/2}	6.6	1.69
40	Zr	(Kr)(5s) ² (4d) ²	³ F ₂	7.0	1.59
41	Nb	(Kr)(5s)(4d) ⁴	⁶ D _{1/2}	6.8	1.59
42	Mo	(Kr)(5s)(4d) ⁵	⁷ S ₃	7.2	1.52
43	Tc	(Kr)(5s) ² (4d) ⁵	⁶ S _{5/2}	Not known	1.39
44	Ru	(Kr)(5s)(4d) ⁷	⁵ F ₅	7.5	1.41
45	Rh	(Kr)(5s)(4d) ⁸	⁴ F _{9/2}	7.7	1.36
46	Pd	(Kr)(4d) ¹⁰	¹ S ₀	8.3	0.57
47	Ag	(Kr)(5s)(4d) ¹⁰	² S _{1/2}	7.6	1.29
48	Cd	(Kr)(5s) ² (4d) ¹⁰	¹ S ₀	9.0	1.18
49	In	(Kr)(5s) ² (4d) ¹⁰ (5p)	² P _{1/2}	5.8	1.38
50	Sn	(Kr)(5s) ² (4d) ¹⁰ (5p) ²	³ P ₀	7.3	1.24
51	Sb	(Kr)(5s) ² (4d) ¹⁰ (5p) ³	⁴ S _{3/2}	8.6	1.19
52	Te	(Kr)(5s) ² (4d) ¹⁰ (5p) ⁴	³ P ₂	9.0	1.11
53	I	(Kr)(5s) ² (4d) ¹⁰ (5p) ⁵	² P _{3/2}	10.4	1.04
54	Xe	(Kr)(5s) ² (4d) ¹⁰ (5p) ⁶	¹ S ₀	12.1	0.99
55	Cs	(Xe)(6s)	² S _{1/2}	3.9	2.52
56	Ba	(Xe)(6s) ²	¹ S ₀	5.2	2.06
57	La	(Xe)(6s) ² (5d)	² D _{3/2}	5.6	1.92
58	Ce	(Xe)(6s) ² (4f)(5d)	³ H ₅	6.9	1.98
59	Pr	(Xe)(6s) ² (4f) ³	⁴ I _{9/2}	5.8	1.94
60	Nd	(Xe)(6s) ² (4f) ⁴	⁵ I ₄	6.3	1.92
61	Pm	(Xe)(6s) ² (4f) ⁵	⁶ H _{5/2}	Not known	1.88
62	Sm	(Xe)(6s) ² (4f) ⁶	⁷ F ₀	5.6	1.84
63	Eu	(Xe)(6s) ² (4f) ⁷	⁸ S _{7/2}	5.7	1.83
64	Gd	(Xe)(6s) ² (4f) ⁷ (5d)	⁹ D ₂	6.2	1.71
65	Tb	(Xe)(6s) ² (4f) ⁹	⁶ H _{15/2}	6.7	1.78
66	Dy	(Xe)(6s) ² (4f) ¹⁰	⁵ I ₈	6.8	1.75
67	Ho	(Xe)(6s) ² (4f) ¹¹	⁴ I _{15/2}	Not known	1.73
68	Er	(Xe)(6s) ² (4f) ¹²	³ H ₆	Not known	1.70
69	Tm	(Xe)(6s) ² (4f) ¹³	² F _{7/2}	Not known	1.68
70	Yb	(Xe)(6s) ² (4f) ¹⁴	¹ S ₀	6.2	1.66
71	Lu	(Xe)(6s) ² (4f) ¹⁴ (5d)	² D _{3/2}	5.0	1.55
72	Hf	(Xe)(6s) ² (4f) ¹⁴ (5d) ²	³ F ₂	5.5	1.48
73	Ta	(Xe)(6s) ² (4f) ¹⁴ (5d) ³	⁴ F _{3/2}	7.9	1.41
74	W	(Xe)(6s) ² (4f) ¹⁴ (5d) ⁴	⁵ D ₀	8.0	1.36
75	Re	(Xe)(6s) ² (4f) ¹⁴ (5d) ⁵	⁶ S _{5/2}	7.9	1.31
76	Os	(Xe)(6s) ² (4f) ¹⁴ (5d) ⁶	⁵ D ₄	8.7	1.27

Periodic Table

Z	Element	Configuration	Term ¹	Ionization Potential eV	Radius ² in Å
77	Ir	(Xe)(6s) ² (4f) ¹⁴ (5d) ⁷	⁴ F _{9/2}	9.2	1.23
78	Pt	(Xe)(6s)(4f) ¹⁴ (5d) ⁹	³ D ₃	9.0	1.22
79	Au	(Xe)(6s)(4f) ¹⁴ (5d) ¹⁰	² S _{1/2}	9.2	1.19
80	Hg	(Xe)(6s) ² (4f) ¹⁴ (5d) ¹⁰	¹ S ₀	10.4	1.13
81	Tl	(Xe)(6s) ² (4f) ¹⁴ (5d) ¹⁰ (6p)	² P _{1/2}	6.1	1.32
82	Pb	(Xe)(6s) ² (4f) ¹⁴ (5d) ¹⁰ (6p) ²	³ P ₀	7.4	1.22
83	Bi	(Xe)(6s) ² (4f) ¹⁴ (5d) ¹⁰ (6p) ³	⁴ S _{3/2}	7.3	1.30
84	Po	(Xe)(6s) ² (4f) ¹⁴ (5d) ¹⁰ (6p) ⁴	³ P ₂	8.4	1.21
85	At	(Xe)(6s) ² (4f) ¹⁴ (5d) ¹⁰ (6p) ⁵	² P _{3/2}	Not known	1.15
86	Rn	(Xe)(6s) ² (4f) ¹⁴ (5d) ¹⁰ (6p) ⁶	¹ S ₀	10.7	1.09
87	Fr	(Rn)(7s)		Not known	2.48
88	Ra	(Rn)(7s) ²	¹ S ₀	5.3	2.04
89	Ac	(Rn)(7s) ² (6d)	² D _{3/2}	6.9	1.90
90	Th	(Rn)(7s) ² (6d) ²	³ F ₂		
91	Pa	(Rn)(7s) ² (5f) ² (6d)	⁴ K _{11/2}		
92	U	(Rn)(7s) ² (5f) ³ (6d)	⁵ L ₆		
93	Np	(Rn)(7s) ² (5f) ⁴ (6d)	⁶ L _{11/2}		
94	Pu	(Rn)(7s) ² (5f) ⁶	⁷ F ₀		
95	Am	(Rn)(7s) ² (5f) ⁷	⁸ S _{7/2}		
96	Cm	(Rn)(7s) ² (5f) ⁷ (6d)	⁹ D ₂		
97	Bk	(Rn)(7s) ² (5f) ⁹	⁶ H _{15/2}		
98	Cf	(Rn)(7s) ² (5f) ¹⁰	⁵ I ₈		
99	Es	(Rn)(7s) ² (5f) ¹¹	⁴ I _{15/2}		
100	Fm	(Rn)(7s) ² (5f) ¹²	³ H ₆		
101	Md	(Rn)(7s) ² (5f) ¹³	² F _{7/2}		
102	No	(Rn)(7s) ² (5f) ¹⁴	¹ S ₀		

¹Term designation is equivalent to spectroscopic description.²Radius is defined by the peak of the calculated charge density of the outermost orbital.

Supplement 14-C

A Brief Discussion of Molecules

The purpose of this supplement is to outline the basic approach to the study of simple molecules. We discuss the H_2 molecule in some detail, so as to provide an understanding of terms like *molecular orbitals* and *valence bonds*. Quantum chemistry has become a field well served by massive computers. Our discussion does not really provide an entry into this field. It is extremely simple-minded, and its only justification is that it provides an insight into the basic mechanisms that lead to molecular binding. Anything more depends on an understanding of electron–electron correlations, and these are way beyond the scope of this book. Our approach will follow the one followed in the discussion of the H_2^+ molecule. It is based on the fact that the nuclei are much more massive than the electrons, and that therefore a good first approximation treats the nuclei as *frozen*, with their location determined by the electronic distribution. We discuss the H_2 molecule as a prototype of other simple diatomic molecules.

The H_2 Molecule

The H_2 molecule is a more complicated system, because there are two electrons present, and the exclusion principle therefore plays a role. As in the case of the H_2^+ molecule, we treat the nuclei as fixed.

The nuclei (protons here) will be labeled A and B , and the two electrons 1 and 2, respectively (Fig. 14C-1). The Hamiltonian has the form

$$H = H_1 + H_2 + \frac{e^2}{4\pi\epsilon_0 r_{12}} + \frac{e^2}{4\pi\epsilon_0 R_{AB}} \quad (14C-1)$$

where

$$H_i = \frac{p_i^2}{2m} - \frac{e^2}{4\pi\epsilon_0 r_{Ai}} - \frac{e^2}{4\pi\epsilon_0 r_{Bi}} \quad (i = 1, 2) \quad (14C-2)$$

depends only on the coordinates of the electron i relative to the nuclei. We will again compute an upper bound to $E(R_{AB})$ by constructing the expectation value of H with a trial wave function. Since

$$\tilde{H}_i = H_i + \frac{e^2}{4\pi\epsilon_0 R_{AB}} \quad (14C-3)$$

are just Hamiltonians for the H_2^+ molecule (14-50) it seems reasonable to take as our trial wave function a product of two functions of the type shown in the first line of (14-51):

$$\psi_g(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2[1 + S(R_{AB})]} [\psi_A(\mathbf{r}_1) + \psi_B(\mathbf{r}_1)][\psi_A(\mathbf{r}_2) + \psi_B(\mathbf{r}_2)]X_{\text{singlet}} \quad (14C-4)$$

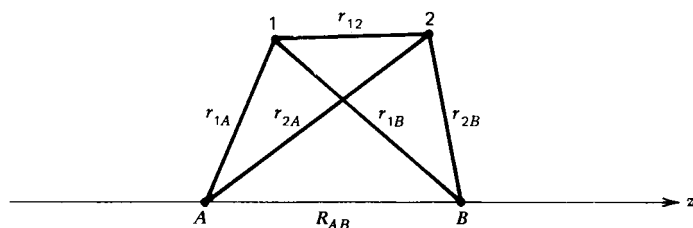


Figure 14C-1 Coordinate labels in the discussion of the H_2 molecule.

The electron spin state is a singlet, since the spatial part of the wave function is taken to be symmetric. In this trial wave function, *each electron is associated with both protons*; that is, the trial wave function is said to be a product of *molecular orbitals*. The description in terms of molecular orbitals is sometimes called the *MO* method.

The calculation of $\langle \psi_g | H | \psi_g \rangle$ yields

$$\begin{aligned}
 & \left\langle \psi_g \left| \left(\tilde{H}_1 - \frac{e^2}{4\pi\epsilon_0 R_{AB}} \right) + \left(\tilde{H}_2 - \frac{e^2}{4\pi\epsilon_0 R_{AB}} \right) + \frac{e^2}{4\pi\epsilon_0 r_{12}} + \frac{e^2}{4\pi\epsilon_0 R_{AB}} \right| \psi_g \right\rangle \\
 &= E(R_{AB}) + E(R_{AB}) + \left\langle \psi_g \left| \frac{e^2}{4\pi\epsilon_0 r_{12}} \right| \psi_g \right\rangle - \frac{e^2}{4\pi\epsilon_0 R_{AB}} \quad (14C-5) \\
 &= 2E(R_{AB}) - \frac{e^2}{4\pi\epsilon_0 R_{AB}} + \left\langle \psi_g \left| \frac{e^2}{4\pi\epsilon_0 r_{12}} \right| \psi_g \right\rangle
 \end{aligned}$$

where $E(R_{AB})$ is the energy of the H_2^+ molecule calculated in (14.56). The first-order electron–electron repulsion contribution can also be calculated, and when the total energy so computed is minimized with respect to the separation R_{AB} , it is found that the binding energy and internuclear separation are given by

$$\begin{aligned}
 E_b &= -2.68 \text{ eV} \\
 R &= 0.85 \text{ \AA}
 \end{aligned}$$

The experimental values are

$$\begin{aligned}
 E_b &= -4.75 \text{ eV} \\
 R &= 0.74 \text{ \AA}
 \end{aligned}$$

Evidently the approximation is not a very good one. We noted in our discussion of the H_2^+ molecule that the trial wave functions (the *MO*'s) are inaccurate for small proton–proton separations, and the fact that the *MO*'s are too spread out in space shows up in the numbers above. The trial wave function also has some undesirable features for large R_{AB} . The product in (14C-4) may be rewritten in the form

$$\begin{aligned}
 & [\psi_A(\mathbf{r}_1) + \psi_B(\mathbf{r}_1)][\psi_A(\mathbf{r}_2) + \psi_B(\mathbf{r}_2)] \\
 &= [\psi_A(\mathbf{r}_1)\psi_A(\mathbf{r}_2) + \psi_B(\mathbf{r}_1)\psi_B(\mathbf{r}_2)] + [\psi_A(\mathbf{r}_1)\psi_B(\mathbf{r}_2) + \psi_A(\mathbf{r}_2)\psi_B(\mathbf{r}_1)] \quad (14C-6)
 \end{aligned}$$

The first term is called an “ionic” term, since it describes both electrons bound to one proton or the other. The second term, the “covalent” term, is a description in terms of linear combinations of atomic orbitals (LCAO). Our trial wave function thus implies, since the two terms enter with equal weight, that for large R_{AB} the molecule is as likely to dissociate into the ions H^+ and H^- as it is into two hydrogen atoms, and this is patently false.

The Valence Bond Method

The last difficulty can be avoided with the use of the *valence bond* (also called Heitler-London) method, in which linear combinations of atomic orbitals are used. The singlet wave function used as a trial wave function in the variational principle is taken to be

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \left\{ \frac{1}{2[1 + S^2(R_{AB})]} \right\}^{1/2} [\psi_A(\mathbf{r}_1)\psi_B(\mathbf{r}_2) + \psi_A(\mathbf{r}_2)\psi_B(\mathbf{r}_1)] X_{\text{singlet}} \quad (14C-7)$$

where, as before, the $\psi_A(\mathbf{r}_i)$ are hydrogenic wave functions for the i th electron about proton A. We could, in principle, add a triplet term to our variational trial wave function. However, a triplet wave function must be spatially antisymmetric and yields low probability for the electrons being located in the region between the protons. We saw in our discussion of the H_2^+ molecule that just this configuration led to the lowest energy. Although it is not immediately obvious that the attraction is still largest in this configuration when there are two electrons that repel each other in the system, it is in fact so. The results of a variational calculation with the VB trial wave function are

$$E_b = -3.14 \text{ eV} \\ R = 0.87 \text{ \AA}$$

This is not a significant improvement over the *MO* results, for the simple reason that the inadequacy of the trial wave functions for small R_{AB} carries more weight. There should be no question about the quantitative successes of quantum mechanics in molecular physics. More sophisticated trial wave functions have to be used; for example, a 50-term trial wave function yields complete agreement with observations for the H_2 molecule, but it does not, as the *MO* and *VB* functions do, give us something of a qualitative feeling of what goes on between the atoms. In what follows, we will explore the relevance of these approaches to a qualitative understanding of some aspects of chemistry.

The expectation value of H for the H_2 molecule in the VB approach has the following schematic form:

$$\begin{aligned} \langle \psi | H | \psi \rangle &= \frac{1}{2(1 + S^2)} \langle \psi_{A1}\psi_{B2} + \psi_{A2}\psi_{B1} | H | \psi_{A1}\psi_{B2} + \psi_{A2}\psi_{B1} \rangle \\ &= \frac{1}{1 + S^2} \left\langle \psi_{A1}\psi_{B2} \left| T_1 + T_2 - \frac{e^2}{4\pi\epsilon_0 r_{A1}} - \frac{e^2}{4\pi\epsilon_0 r_{A2}} - \frac{e^2}{4\pi\epsilon_0 r_{B1}} - \frac{e^2}{4\pi\epsilon_0 r_{B2}} \right. \right. \\ &\quad \left. \left. + \frac{e^2}{4\pi\epsilon_0 r_{12}} + \frac{e^2}{4\pi\epsilon_0 R_{AB}} \right| \psi_{A1}\psi_{B2} + \psi_{A2}\psi_{B1} \right\rangle \end{aligned}$$

where T_i is the kinetic energy of the i th electron, and since

$$\begin{aligned} \left(T_1 - \frac{e^2}{4\pi\epsilon_0 r_{A1}} \right) \psi_{A1} &= E_1 \psi_{A1} \\ \langle \psi | H | \psi \rangle &= \frac{1}{2(1 + S^2)} \langle \psi_{A1}\psi_{B2} + \psi_{A2}\psi_{B1} | H | \psi_{A1}\psi_{B2} + \psi_{A2}\psi_{B1} \rangle \\ &= \frac{1}{1 + S^2} \left\langle \psi_{A1}\psi_{B2} \left| 2E_1 - \frac{e^2}{4\pi\epsilon_0 r_{B1}} - \frac{e^2}{4\pi\epsilon_0 r_{A2}} + \frac{e^2}{4\pi\epsilon_0 r_{12}} + \frac{e^2}{4\pi\epsilon_0 R_{AB}} \right| \psi_{A1}\psi_{B2} \right\rangle \\ &\quad + \left\langle \psi_{A1}\psi_{B2} \left| 2E_1 - \frac{e^2}{4\pi\epsilon_0 r_{B2}} - \frac{e^2}{4\pi\epsilon_0 r_{A1}} + \frac{e^2}{4\pi\epsilon_0 r_{12}} + \frac{e^2}{4\pi\epsilon_0 R_{AB}} \right| \psi_{A2}\psi_{B1} \right\rangle \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{1+S^2} \left\{ \left(2E_1 + \frac{e^2}{4\pi\epsilon_0 r_{AB}} \right) (1+S^2) - 2 \frac{e^2}{4\pi\epsilon_0} \left\langle \psi_{A1} \left| \frac{1}{r_{B1}} \right| \psi_{A1} \right\rangle \right. \\
&\quad - 2 \frac{e^2}{4\pi\epsilon_0} S \left\langle \psi_{A1} \left| \frac{1}{r_{A1}} \right| \psi_{B1} \right\rangle + \frac{e^2}{4\pi\epsilon_0} \int \int \frac{|\psi_{A1}|^2 |\psi_{B2}|^2}{r_{12}} \\
&\quad \left. + \int \int \frac{\psi_{A1}^* \psi_{B1} \psi_{B2}^* \psi_{A2}}{r_{12}} \right\} \quad (14C-8)
\end{aligned}$$

In obtaining this, liberal use has been made of symmetry. The terms that can make this expression more negative are

$$\left\langle \psi_{A1} \left| \frac{1}{r_{B1}} \right| \psi_{A1} \right\rangle \quad \text{and} \quad \frac{S}{1+S^2} \left\langle \psi_{A1} \left| \frac{1}{r_{A1}} \right| \psi_{B1} \right\rangle$$

The former is just the attraction of the electron cloud about one proton to the other proton; the second is the overlap of the two electrons (weighted with $1/r_{A1}$). If this can be large, there will be binding. The two electrons can only overlap significantly, however, if their spins are antiparallel; this is a consequence of the exclusion principle. The region of overlap is between the two nuclei, and there the attraction to the nuclei generally overcomes the electrostatic repulsion between the electrons.

In the *MO* picture, too, it is an overlap term—the result in (14-60)—that is crucial to bonding, and again, bonding occurs because the electron charge distribution is large between the nuclei. Thus, although here the orbitals belong to the whole molecule rather than to individual atoms, the physical reason for bonding is the same.

We note that in general there may be several bound states of the nuclei, corresponding to different electronic configurations. For example, if in (14C-7) we take the $\psi(\mathbf{r}_2)$ wave function to be a u_{200} eigenfunction, while the $\psi(\mathbf{r}_1)$ remains a u_{100} eigenfunction, the overlap may be such as to provide a second, more weakly bound state of the protons. We are not going to pursue this, except to point out the important fact that the $E(R)$ is *different for each electronic state*.

The Importance of Unpaired Valence Electrons

An important simplification in the study of electronic charge distributions in molecules occurs because we really do not need to take all electrons into account. In the construction of orbitals, be it valence or molecular, only the outermost electrons, not in closed shells—that is, the so-called *valence electrons*—have a chance to contribute to the bonding. The inner electrons, being closer to the nucleus, are less affected by the presence of another atom in the vicinity.¹ Furthermore, not all valence electrons contribute equally: if two electrons are in a spin 0 state—we call them *paired electrons*—they will *not give rise to bonding*. To see why this is so, consider what happens when an atom with a single valence electron is brought near an atom with two paired electrons. There are two cases to be considered (Fig. 14C-2).

- (a) If the two electrons that are parallel exchange [i.e., are put into a form such as (14C-7) with a minus sign between the terms], then they must be in a triplet

¹It may happen in atoms that even the valence electrons are rather close to the nucleus. This is the case for the rare earths. A consequence of the fact that the outer electrons in 5*d* and 4*f* shells lie close in is that the rare earths are chemically less active than the transition metals ($Z \approx 20-30$).

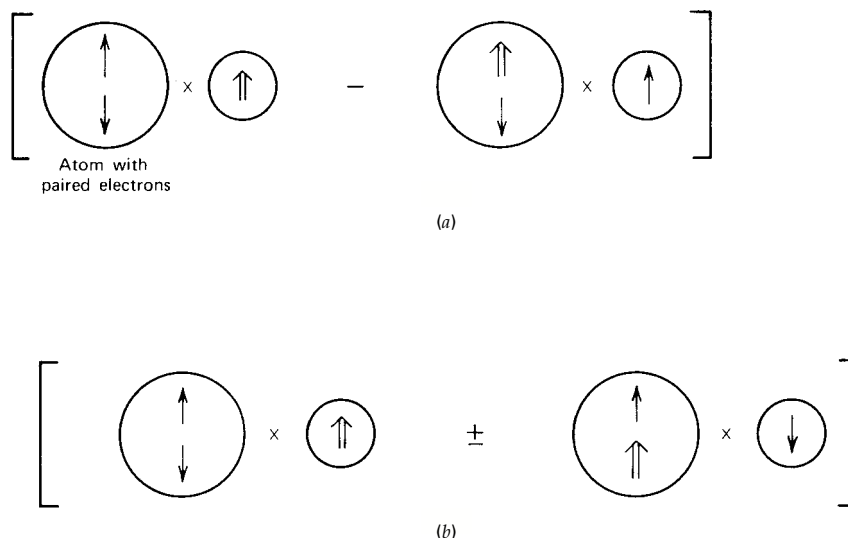


Figure 14C-2 Illustration of why paired electrons do not give rise to bonding. (a) If parallel electrons exchange, the wave function is spatially antisymmetric. (b) If antiparallel electrons exchange, one term in the wave function has electrons in the same spin state, which may require promotion to a higher energy orbit.

state, and hence the spatial wave function of this pair must be antisymmetric. This reduces the overlap, and it turns out that the exchange integral gives a repulsive contribution to the energy.

- (b) When the two electrons that are antiparallel exchange, then one atom finds itself some of the time with two electrons in the same spin state. The original atomic state will frequently no longer be a possible one, and one of the electrons will have to be promoted into another atomic orbital. Sometimes this may cost very little energy, but usually this is not the case, and again bonding is not achieved. *Chemical activity depends on the presence of unpaired outer electrons.* An example of this is the nonexistence of the H-He molecule. In He we have two electrons in the 1s state; promotion of one of them into a 2s state costs a lot of energy. It is for this reason that the atoms for which the outer shells are closed are *inert*. Not all unpaired electrons are of equal significance. As noted before, the unpaired *d*- and *f*-electrons in the transition elements tend to be close to the nucleus, and hence inactive. Thus, mainly *s*- and *p*-electrons in the outer shells contribute to chemical activity. The pairing effect is also responsible for what is called the “saturation of chemical binding forces.” Once two unpaired electrons from different atoms form a singlet state (and cause bonding), they become paired; an electron from a third atom must find an unpaired electron elsewhere—that is, participate in a different bond. Another consequence is that molecules have spin 0 in most cases.

Chapter 15

Time-Dependent Perturbation Theory

15-1 FORMALISM

In previous chapters we were primarily concerned with the stationary states of atomic systems. Much of our information about atoms, molecules, nuclei, and elementary particles comes from studies of transitions between energy levels or decays. Although most of our interest will focus on electromagnetic transitions, which will be treated in separate chapters, other transitions are also of interest. All of these processes involve time-dependent perturbations, and we shall develop the formalism to deal with them.

As before, we start with an unperturbed system, described by a Hamiltonian H_0 . We assume that we know the complete set of unperturbed eigenstates that satisfy

$$H_0|\phi_n\rangle = E_n^{(0)}|\phi_n\rangle \quad (15-1)$$

The kets $|\phi_n\rangle$ are taken to be time independent. The labeling may seem redundant: The n -label describes all the quantum numbers that characterize the states. The ϕ part is just there to remind us that these are eigenstates of the unperturbed Hamiltonian H_0 . We will be interested in solving the Schrödinger equation

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = (H_0 + \lambda V(t))|\psi(t)\rangle \quad (15-2)$$

The fact that our states depend on time means that we are working in the Schrödinger picture. We now apply the procedure we used for time-independent perturbation theory. We expand the $|\psi(t)\rangle$ in terms of the complete set of eigenstate $|\phi_n\rangle$, except that we explicitly insert the time dependence appropriate to these. We write

$$|\psi(t)\rangle = \sum_n c_n(t) e^{-iE_n^{(0)}t/\hbar} |\phi_n\rangle \quad (15-3)$$

The time-dependence is put in so that in the absence of $V(t)$ the $c_n(t)$ would be constants. If we now substitute this expansion into (15-2) we get

$$\sum_n \left[i\hbar \frac{dc_n(t)}{dt} + E_n^{(0)}c_n(t) \right] e^{-iE_n^{(0)}t/\hbar} |\phi_n\rangle = H|\psi(t)\rangle = \sum_n [E_n^{(0)} + \lambda V(t)]c_n(t) e^{-iE_n^{(0)}t/\hbar} |\phi_n\rangle \quad (15-4)$$

that is,

$$\sum_n \left[i\hbar \frac{dc_n(t)}{dt} - \lambda V(t)c_n(t) \right] e^{-iE_n^{(0)}t/\hbar} |\phi_n\rangle = 0 \quad (15-5)$$

We now take the scalar product with $\langle\phi_m|$, use the orthonormality of the unperturbed eigenstates $\langle\phi_m|\phi_n\rangle = \delta_{mn}$, and divide out $e^{-iE_m^{(0)}t/\hbar}$. This finally yields the infinite set of equations

$$i\hbar \frac{dc_m(t)}{dt} = \lambda \sum_n c_n(t) e^{i(E_m^{(0)} - E_n^{(0)})t/\hbar} \langle\phi_m|V(t)|\phi_n\rangle \quad (15-6)$$

We first try to solve this to first order in the parameter λ . As an initial condition at $t = 0$, we take the system to be in a particular state $|\phi_k\rangle$ so that $|\psi(0)\rangle = |\phi_k\rangle$ and thus

$$c_n(0) = \delta_{nk} \quad (15-7)$$

Sometimes the initial state is specified in the distant past, as in a scattering process. In that case (15-7) would read

$$\lim_{t \rightarrow -\infty} c_n(t) = \delta_{nk} \quad (15-8)$$

Departures from this at a later time depend on $\lambda V(t)$. To first order in λ it is enough to substitute (15-7) into the r.h.s. of (15-6). This yields the differential equation (for $m \neq k$)

$$i\hbar \frac{dc_m(t)}{dt} = \lambda e^{i(E_m^{(0)} - E_k^{(0)})t/\hbar} \langle\phi_m|V(t)|\phi_k\rangle \quad (15-9)$$

whose solution is

$$c_m(t) = \frac{\lambda}{i\hbar} \int_0^t dt' e^{i\omega_{mk}t'} \langle\phi_m|V(t')|\phi_k\rangle \quad (15-10)$$

with $\omega_{mk} = (E_m^{(0)} - E_k^{(0)})/\hbar$. One can iterate this and obtain an expression to second order in λ , but the derivation is messy, and we do not need it in this book.¹

The probability that at a later time t a measurement of the state $|\psi(t)\rangle$ yields the quantum numbers of the state $|\phi_n\rangle$ is, according to the expansion postulate,

$$P_n(t) = |\langle\phi_n|\psi(t)\rangle|^2 = |c_n(t)|^2 \quad (15-11)$$

Note that

$$\sum_n P_n(t) = \sum_n |\langle\phi_n|\psi(t)\rangle|^2 = \sum_n \langle\psi(t)|\phi_n\rangle \langle\phi_n|\psi(t)\rangle = \langle\psi(t)|\psi(t)\rangle = 1$$

when account is taken of the completeness of the set of eigenstates $|\phi_n\rangle$ and the fact that the $|\psi(t)\rangle$ are properly normalized to unity.

EXAMPLE 15-1

A particle of charge q in a one-dimensional harmonic oscillator of characteristic frequency ω is placed in an electric field that is turned on and off so that the potential energy is

$$\lambda V(t) = qExe^{-t^2/\tau^2}$$

If the particle is initially in the ground state, what is the probability that after time t , such that $t \gg \tau$, the particle is found in the first excited state of the harmonic oscillator? What is the probability that it is found in the second excited state?

¹See Supplement 15-B [www.wiley.com/college/gasiorowicz] on the *interaction picture* for a neater (but still messy) approach to this subject.

SOLUTION If we label the harmonic oscillator eigenstates by $|n\rangle$, with energy eigenvalues $\hbar\omega(n + 1/2)$, then our formula leads to

$$c_n(t) = \frac{qE}{i\hbar} \int_{-\infty}^t dt' e^{i\omega_0 t'} \langle n|x|0\rangle e^{-i'^2 \tau^2}$$

Since $t \gg \tau$, we may take the upper and lower limits in the time integral in (15-10) to be $\pm\infty$. We therefore get

$$\begin{aligned} c_n(\infty) &= \frac{qE}{i\hbar} \langle n|x|0\rangle \int_{-\infty}^{\infty} dt' e^{-i\omega_0 t'} e^{-i'^2 \tau^2} \\ &= \frac{qE}{i\hbar} \langle n|x|0\rangle \int_{-\infty}^{\infty} dt' e^{-i\omega_0 t'} e^{-i'^2 \tau^2} \\ &= \sqrt{\pi} \frac{qE\tau}{i\hbar} \langle n|x|0\rangle e^{-n^2 \omega^2 \tau^2 / 4} \end{aligned}$$

This means that

$$P_n(\infty) = \pi \frac{q^2 E^2 \tau^2}{\hbar^2} |\langle n|x|0\rangle|^2 e^{-n^2 \omega^2 \tau^2 / 2}$$

To calculate $\langle n|x|0\rangle$ we use (6-35), to get

$$x = \sqrt{\frac{\hbar}{2m\omega}} (A + A^\dagger)$$

and recall that $A|0\rangle = 0$ and $A^\dagger|0\rangle = |1\rangle$. Thus $|\langle n|x|0\rangle|^2 = \frac{\hbar}{2m\omega} \delta_{n1}$ and therefore

$$P_1 = \frac{\pi}{2} \frac{q^2 E^2 \tau^2}{m\hbar\omega} e^{-\omega^2 \tau^2 / 2}$$

We also see that $P_n = 0$ for $n = 2, 3, \dots$

Note the following

1. As $\tau \rightarrow \infty$ $P_1 \rightarrow 0$. When the electric field is turned on very slowly, then the transition probability goes to zero; that is, the system adjusts *adiabatically* to the presence of the field, without being “jolted” into making a transition.
2. The potential, through its x -dependence, will allow only certain transitions to take place. In other words, we see that there is a *selection rule* that operates here. If the potential were proportional to x^2 , for example, then transitions to $n = 2$ would be allowed

15-2 HARMONIC TIME-VARIATION OF THE POTENTIAL

In a large number of examples, the potential has the time dependence given by

$$V(t) = M e^{\mp i\omega t} \quad (15-12)$$

where M is an operator that does not have an explicit time dependence. Such a time dependence may arise either as a result of the system being placed in an external potential that has a harmonic time dependence (for example an atom irradiated by a laser beam) or because the potential describes an emission or absorption of a particle in the $k \rightarrow m$ transi-

tion. In that case, the $\pm\omega$ corresponds to the emission/absorption of a particle. For this time dependence,

$$c_m(t) = \frac{\lambda}{i\hbar} \langle \phi_m | M | \phi_k \rangle \int_0^t dt' e^{i\omega_{mk}t'} e^{\mp i\omega t'}$$

The integral is easily done:

$$\int_0^t dt' e^{i\omega_{mk}t'} e^{\mp i\omega t'} = \frac{e^{i(\omega_{mk} \mp \omega)t} - 1}{i(\omega_{mk} \mp \omega)} = e^{i(\omega_{mk} \mp \omega)t/2} \frac{\sin(\omega_{mk} \mp \omega)t/2}{(\omega_{mk} \mp \omega)/2}$$

The absolute square of $c_m(t)$ then yields

$$\frac{\lambda^2}{\hbar^2} |\langle \phi_m | M | \phi_k \rangle|^2 \left(\frac{\sin \frac{\omega_{mk} \mp \omega}{2} t}{\frac{\omega_{mk} \mp \omega}{2}} \right)^2 \quad (15-13)$$

The sinusoidal variation is characteristic of the transition rate for an atom in a laser beam. We will discuss this matter in detail in Chapter 18.

For transitions involving the emission or absorption of a particle with energy $E = \hbar\omega$, we are interested in the long-time behavior of the transition rate. Let us look at the form

$$F(t) = \frac{4}{\Delta^2} \sin^2 \frac{t\Delta}{2}$$

with

$$\Delta = \frac{E_m^{(0)} - E_k^{(0)} \mp \hbar\omega}{\hbar}$$

Figure 15-1 shows the behavior of this function. For large t it becomes strongly peaked at $\Delta = 0$, and away from $\Delta = 0$ it oscillates very rapidly. This is the kind of behavior that

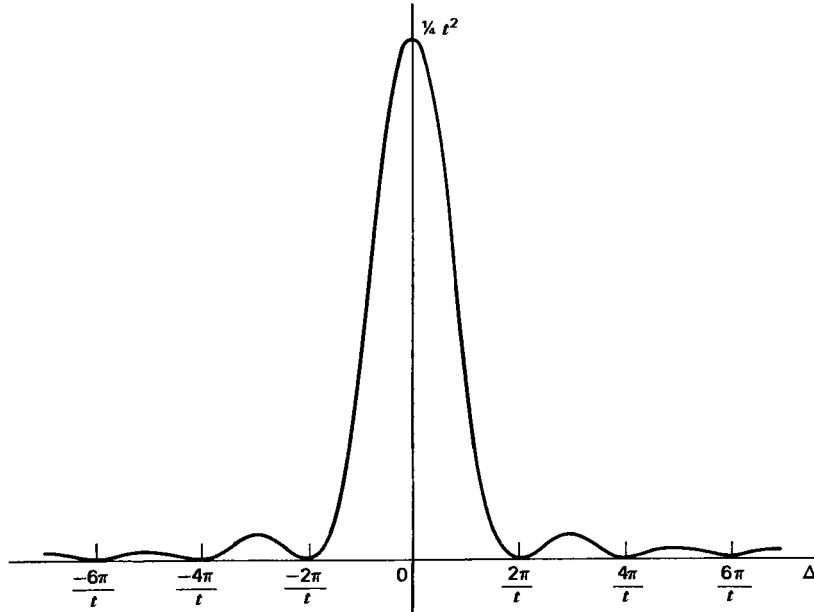


Figure 15-1 Plot of the function $1/\Delta^2 \sin^2 t\Delta/2$ versus Δ

we associate with a delta function. In fact, if $f(\Delta)$ is a smooth function of Δ , then, for t large,

$$\begin{aligned} \int_{-\infty}^{\infty} f(\Delta) \frac{4}{\Delta^2} \sin^2 \frac{t\Delta}{2} d\Delta &\approx f(0) \int_{-\infty}^{\infty} d\Delta \frac{4}{\Delta^2} \sin^2 \frac{t\Delta}{2} \\ &= 2tf(0) \int_{-\infty}^{\infty} dy \frac{1}{y^2} \sin^2 y = 2\pi tf(0) \end{aligned}$$

that is, for t large,

$$\frac{4}{\Delta^2} \sin^2 \frac{t\Delta}{2} \rightarrow 2\pi t \delta(\Delta) = 2\pi \hbar t \delta(E_m^{(0)} - E_k^{(0)} \mp \hbar\omega) \quad (15-14)$$

Thus the transition probability in (15-13) after a long time grows linearly with t , and the *transition probability per unit time* is

$$\Gamma_{k \rightarrow m} = \frac{2\pi}{\hbar} |\langle \phi_m | \lambda M | \phi_k \rangle|^2 \delta(E_m^{(0)} - E_k^{(0)} \mp \hbar\omega) \quad (15-15)$$

The delta function shows that transitions are only induced if

$$\hbar\omega = |E_m^{(0)} - E_k^{(0)}|$$

This is a statement of energy conservation. If $E_k^{(0)} < E_m^{(0)}$ then the system is excited to a higher energy state, and this is only possible if energy is supplied by the potential $Me^{\mp i\omega t}$. Note that [See eq. (5-49)]

$$\left\langle \frac{dH}{dt} \right\rangle = \left\langle \frac{\partial H}{\partial t} \right\rangle$$

so that with a time-dependent potential, energy is not conserved. Here we see explicitly how much energy a time-dependent potential can absorb or emit.

The result needs some discussion. For fixed, precise ω , (15-13) gives a perfectly sensible oscillating result. The probability for a transition oscillates with frequency $\omega_{mk} \mp \omega$. The reason we look at large times $t \rightarrow \infty$ is that we are interested in transition rates for atoms that emit a photon or for radioactive decays such as α -decay that occur in nuclei. In such cases the reader may well feel swindled. First, the manipulations leading to (15-15) are suspect. They involve limits such as “ t large,” which are vague, and which clearly cannot hold for all large t , since a transition probability that grows linearly with time must sooner or later exceed unity. Second, the result for a perfectly reasonable quantity such as a transition probability per unit time—that is, a *transition rate*—should not include a delta function that is not a sensible smooth function.

The first difficulty is addressed in Supplement 15-A [www.wiley.com/college/gasiorowicz]. There we discuss a more accurate treatment of the decay probability, which gives rise to the familiar exponential decay rate, whose outcome is the formula

$$N(t) = N(0)e^{-\Gamma t} \quad (15-16)$$

The second difficulty involves the problem of the delta function. We discuss it in the framework of two explicit examples: (1) the decay of an excited state of an atom, described by $|\phi_k\rangle$ to another state, $|\phi_m\rangle$, accompanied by the emission of a photon of energy $\hbar\omega$, and (2) a nuclear decay in which a nucleus described by $|A\rangle$ decays to a state $|B\rangle$, with the emission of a neutron. The neutron’s energy, because of the conservation law, must have the energy $E = E_A - E_B$. This will manifest itself in the presence of $\delta(E_A - E_B - E)$. Never-

theless, that is not the whole story. In each case, the energy of the emitted particle does not specify the state of the emitted particle *uniquely*. We are always interested in the detection of the final particle (photon or neutron, here) in some momentum range. If we are interested in the momentum range $(\mathbf{p}, \mathbf{p} + \Delta\mathbf{p})$, then what we really want to calculate is

$$\Gamma = \sum_{\Delta\mathbf{p}} \Gamma_{k \rightarrow m}(\mathbf{p}) \quad (15-17)$$

where the sum is over all states in that range. It is this sum that corresponds to an integration over the delta function, removes the infinite spike, and evaluates some function at the location of the spike. To see this explicitly, we work this out for an arbitrary emitted particle.

15-3 PHASE SPACE

Consider the emission of a particle, photon, or neutron. Far away the particle will have to be described by a plane-wave of the form

$$\psi(\mathbf{r}, t) = \frac{1}{\sqrt{V}} e^{i(\mathbf{p}\cdot\mathbf{r} - Et)/\hbar} \quad (15-18)$$

Here V is the volume of the box in which the process takes place. This factor is there to normalize the wave function. We take the box to be cubical, of side L and very large. Under those circumstances the shape of the box does not matter. The specification of the momentum in the box is most easily obtained if we use periodic boundary conditions, so that

$$\psi(x + L, y, z) = \psi(x, y, z) \quad (15-19)$$

and so on. This implies that

$$e^{ip_x L/\hbar} = e^{ip_y L/\hbar} = e^{ip_z L/\hbar} = 1$$

The momenta are therefore specified by

$$p_x = \frac{2\pi\hbar}{L} n_1; \quad p_y = \frac{2\pi\hbar}{L} n_2; \quad p_z = \frac{2\pi\hbar}{L} n_3 \quad (15-20)$$

where n_1, n_2 , and n_3 are integers, positive and negative. Thus the sum over an interval $\Delta\mathbf{p}$ becomes a sum over integers

$$\Gamma = \sum_{\Delta n_1} \sum_{\Delta n_2} \sum_{\Delta n_3} \Gamma_{k \rightarrow m}(\mathbf{n}) \rightarrow \int d^3n \Gamma_{k \rightarrow m}(\mathbf{n}) \quad (15-21)$$

The sum becomes an integral because in a large volume the triplet of integers $\mathbf{n}$ consists of very large numbers, so that Δn_i becomes infinitesimal. Next, we note that

$$d^3n = dn_1 dn_2 dn_3 = \left(\frac{L}{2\pi\hbar}\right)^3 dp_x dp_y dp_z = \frac{V}{(2\pi\hbar)^3} d^3p \quad (15-22)$$

The sum in (15-21) therefore leads to

$$\Gamma_{k \rightarrow m} = \frac{2\pi}{\hbar} \int \frac{V d^3p}{(2\pi\hbar)^3} |\langle \phi_m | \lambda M | \phi_k \rangle|^2 \delta(E_m^{(0)} - E_k^{(0)} + E) \quad (15-23)$$

We have chosen a definite sign for the energy E , since we are discussing emission.

To evaluate this further, we take the volume element in $\mathbf{p}$ -space and write it in terms of spherical (momentum-space) coordinates. Thus

$$d^3p = d\Omega_{\mathbf{p}} p^2 dp \quad (15-24)$$

Here the factor $d\Omega_{\mathbf{p}} = \sin \theta_{\mathbf{p}} d\theta_{\mathbf{p}} d\phi_{\mathbf{p}}$ is the infinitesimal solid angle about the direction of $\mathbf{p}$. For the rest, we are left with

$$\begin{aligned} \Gamma_{k \rightarrow m} &= \frac{2\pi}{\hbar} \int d\Omega_{\mathbf{p}} \frac{V}{(2\pi\hbar)^3} \int p^2 dp |\langle \phi_m | \lambda M | \phi_k \rangle|^2 \delta(E_m^{(0)} + E - E_k^{(0)}) \\ &= \frac{2\pi}{\hbar} \int d\Omega_{\mathbf{p}} \frac{V}{(2\pi\hbar)^3} \int dE \left(\frac{p^2 dp}{dE} \right) |\langle \phi_m | \lambda M | \phi_k \rangle|^2 \delta(E_m^{(0)} + E - E_k^{(0)}) \quad (15-25) \\ &= \frac{2\pi}{\hbar} \int d\Omega_{\mathbf{p}} \frac{V}{(2\pi\hbar)^3} \left(\frac{p^2 dp}{dE} |\langle \phi_m | \lambda M | \phi_k \rangle|^2 \right)_{E=E_k^{(0)}-E_m^{(0)}} \end{aligned}$$

In (15-25) the integral is over whatever solid angle we want to include. This may just be the solid angle subtended by a fixed detector; if we don't care in which direction the outgoing particle goes, we integrate over the whole sphere and get 4π . The factor in the parentheses is to be evaluated at the value of E , specified by the energy-conservation delta function. Note that the volume factor V in the above expression is welcome; it just cancels the square of $1/\sqrt{V}$ that appears in the wave function of the outgoing particle.

If the outgoing particle is a photon, then the relation between energy and momentum is $E = pc$, so that the phase space factor is

$$\frac{p^2 dp}{dE} = \frac{p^2}{c} = \frac{E^2}{c^3} = \frac{\hbar^2}{c^3} \omega^2 \quad (15-26)$$

where we have used the relation $E = \hbar\omega$. If the outgoing particle has energy given by $E = \frac{p^2}{2m}$, then

$$\frac{p^2 dp}{dE} = p^2 \frac{m}{p} = mp = m\sqrt{2mE} \quad (15-27)$$

We may rewrite the factor that counts the number of states as follows:

$$\frac{V}{(2\pi\hbar)^3} d\Omega_{\mathbf{p}} \frac{p^2 dp}{dE} = \frac{V d^3p}{(2\pi\hbar)^3 dE} = \frac{d^3n}{dE} \equiv \rho(E) \quad (15-28)$$

The factor is the number of states in the energy interval dE and is therefore called the *density of states*. We may therefore write the transition rate in the form

$$\Gamma_{k \rightarrow m} = \frac{2\pi}{\hbar} |\langle \phi_m | \lambda M | \phi_k \rangle|^2 \rho(E) \quad (15-29)$$

If we have several free particles in the final state, (15-23) generalizes to

$$\Gamma_{i \rightarrow f} = \frac{2\pi}{\hbar} \int \prod_{\text{indep. momenta}} \frac{V d^3p_k}{(2\pi\hbar)^3} |\langle \phi_f | \lambda M | \phi_i \rangle|^2 \delta\left(E_f^{(0)} + \sum_k E_k - E_i^{(0)}\right) \quad (15-30)$$

The delta function again expresses energy conservation. Here all of the energy carried off by the free particles is equal to the change of energy in the system, and *the integration is over independent momenta*. Thus if a massive unstable article decays into three particles, there are only two independent momenta, since the third is fixed by momentum conservation.

Note, however, that the product of the phase space factors in (15-30) is over *all* the particles in the final state. It therefore involves a factor V^N if there are N particles in the final state. This is needed to cancel the square of the $1/\sqrt{V}$ factors for the N particles. Equivalently, we could have written (15-30) as an integral over *all* momenta, with an additional delta function that includes a statement of momentum conservation. The rate is then given by

$$\begin{aligned} \Gamma_{i \rightarrow f} &= \frac{2\pi}{\hbar} \int \int \cdots \int \prod_k \frac{V d^3 p_k}{(2\pi\hbar)^3} |\langle \phi_f | \lambda M | \phi_i \rangle|^2 \delta(E_f^{(0)} + \sum_k E_k - E_i^{(0)}) \delta(\mathbf{p}_f - \sum_k \mathbf{p}_k - \mathbf{p}_i) \end{aligned} \quad (15-31)$$

In the case of atomic transitions, the mass of the atom is so much larger than the (photon energy/ c^2) that we may treat the nucleus fixed in space. Under these circumstances, momentum conservation becomes irrelevant.

The number $\Gamma_{i \rightarrow f}$ represents the probability for the transition $i \rightarrow f$ divided by the time during which the perturbation acts. This time must be long so that the transition probability is proportional to t , but it cannot be too long. If we ask for the probability that the initial state remain intact, we get

$$P_i(t) = 1 - t \sum_{f \neq i} \Gamma_{i \rightarrow f}$$

where the sum is over all accessible states—that is, states not forbidden by energy conservation or selection rules. This expression ceases to make sense for long enough times, since all probabilities must be positive. As indicated above, we show in Supplement 15-A [www.wiley.com/college/gasiorowicz] that the above just form the first two terms of the expansion of

$$P_i(t) = e^{-\Gamma t} \quad (15-32)$$

where $\Gamma = \sum_{f \neq i} \Gamma_{i \rightarrow f}$. From the above formula we may derive a *lifetime* for the initial state, defined by

$$\tau = \frac{1}{\Gamma} \quad (15-33)$$

Given that the initial state has a finite lifetime, one would expect that its energy would be indeterminate, limited by the uncertainty principle for energy and time. We thus would expect its energy to be uncertain by $\Delta E \geq \hbar/\tau$. In atomic transitions, for example, this is seen when the intensity of the spectral lines corresponding to the emitted photon frequency ω_0 is not completely sharp, but is given by

$$I(\omega) = \frac{\Gamma/2}{(\omega - \omega_0)^2 + \Gamma^2/4} \quad (15-34)$$

The width of the line is Γ , and this is a measure of the uncertainty of the energy. The line shape is called a Lorentzian line shape. In the limit that $\Gamma \rightarrow 0$, we get, as a consequence of the formula

$$\lim_{\eta \rightarrow 0} \frac{\eta}{(\omega - \omega_0)^2 + \eta^2} = \pi \delta(\omega - \omega_0) \quad (15-35)$$

the line shape represented by the energy conservation delta function. We will apply the results obtained in this chapter to the discussion of atomic transitions in hydrogen in Chapter 17.

PROBLEMS

1. A hydrogen atom is placed in a time-dependent electric field pointing in the z -direction, with magnitude $E(t)$ given by

$$E(t) = 0 \quad t < 0 \\ = E_0 e^{-\tau} \quad t > 0$$

What is the probability that as $t \rightarrow \infty$ the hydrogen atom, if initially in the ground state, makes a transition to the $2p$ state?

2. Consider a particle in an infinite well, with $V(x) = 0$ for $0 \leq x \leq a$ and $V(x) = \infty$ everywhere else. The potential in the range $0 \leq x \leq a$ changes by an additional term

$$V_1(x) = \lambda \left(x - \frac{a}{2} \right) \sin \omega t$$

- (a) Calculate the probability that a particle in the ground state ($n = 1$) makes a transition to the first excited state ($n = 2$).
 (b) What is the probability that it makes a transition to the second excited state ($n = 3$)?
 (c) What happens to these results as $\omega \rightarrow 0$?
3. Repeat the above calculation with $\sin \omega t$ replaced by $e^{-t/\tau}$. In part (c), consider the limit $\tau \rightarrow \infty$. Note that in both cases (c) shows that for a very slowly varying perturbation transitions become strongly suppressed. See Problem 6.
4. Consider a particle in the n th state of a one-dimensional harmonic oscillator, whose spectrum is $E_n = \hbar\omega(n + \frac{1}{2})$. Suppose the system is perturbed by

$$V(t) = 0 \quad t < 0 \\ = \lambda x \cos \omega_1 t e^{-\alpha t} \quad t > 0$$

Calculate the probability of a transition to the m th state. For what values of m are the transitions allowed? Discuss any special features that arise when $\omega_1 \rightarrow \omega$ and/or $\alpha \rightarrow 0$. (*Hint:* Use the creation and annihilation operator method to calculate the matrix element.)

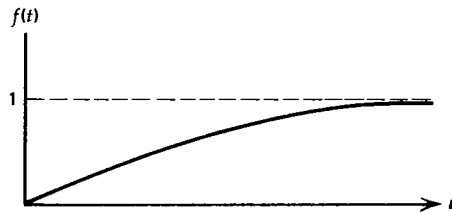
5. Suppose a particle of rest mass M decays into two particles of rest mass m_1 and m_2 , respectively. Use the relativistic relation between energy and momentum to compute the density of states ρ that appears in (15-29). [*Hint:* There is only one independent momentum—say, $\mathbf{p}$ —and what is needed is

$$\int \frac{d^3\mathbf{p}}{(2\pi\hbar)^6} \delta \left(E_{\text{initial}} - \sum_{\text{final states}} E \right)$$

6. In this problem the *adiabatic theorem* is to be illustrated. The theorem states that if the Hamiltonian is changed very slowly from H_0 to H , then a system in a given eigenstate of H_0 goes over into the corresponding eigenstate of H , but does not make any transitions. To be specific, consider the ground state, so that

$$H_0 \phi_0 = E_0 \phi_0$$

Let $V(t) = f(t) V$, where $f(t)$ is a slowly varying function, as shown in the graph.



If the ground state of $H = H_0 + V$ is $|w_0\rangle$, the theorem states that

$$|\langle w_0 | \psi(t) \rangle| \rightarrow 1$$

The steps to be carried out are the following:

(a) Show that

$$\frac{1}{i\hbar} \int_0^t dt' e^{i(E_m^0 - E_0^0)t'/\hbar} f(t') \rightarrow \frac{e^{i(E_m^0 - E_0^0)t/\hbar}}{E_m^0 - E_0^0}$$

for times t such that $f(t) = 1$. Use the fact that

$$\frac{df(t')}{dt'} \ll \frac{E_m^0 - E_0^0}{\hbar} f(t')$$

Either construct an example of a function $f(t)$ or use integration by parts; that is, write

$$e^{i\omega t'} = \frac{1}{i\omega} \frac{d}{dt'} e^{i\omega t'}$$

in the preceding.

(b) Calculate $\psi(t)$ using (15-3) and (15-10). Compare this with the formula (11-15), which here reads

$$|w_0\rangle = |\phi_0\rangle + \sum_{m \neq 0} \frac{\langle \phi_m | V | \phi_0 \rangle}{E_0^{(0)} - E_m^{(0)}} |\phi_m\rangle$$

and thus show that

$$|\langle w_0 | \psi(t) \rangle| \rightarrow 1$$

7. Nuclei sometimes decay from excited states to the ground state by *internal conversion*, a process in which one of the $1s$ electrons is emitted instead of a photon. Let the initial and final nuclear wave functions be

$$\phi_I(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_A) \quad \text{and} \quad \phi_F(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_A)$$

where $\mathbf{r}_i$ ($i = 1, 2, \dots, Z$) describe the protons and $\mathbf{r}_{Z+1} \dots \mathbf{r}_A$ the neutrons. The perturbation giving rise to the transition is just the nucleus-electron interaction

$$V = - \sum_{i=1}^Z \frac{e^2}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_i|}$$

where $\mathbf{r}$ is the electron coordinate. Thus the matrix element is given by

$$- \int d^3\mathbf{r} \int d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_A \phi_F^* \frac{e^{-i\mathbf{p}\cdot\mathbf{r}\hbar}}{\sqrt{V}} \sum_{i=1}^Z \frac{e^2}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_i|} \phi_I \psi_{100}(\mathbf{r})$$

- (a) What is the magnitude of $\mathbf{p}$, the free electron momentum?
 (b) Calculate the rate for the process for a dipole transition in terms of

$$\mathbf{d} = \sum \int d^3\mathbf{r}_1 \cdots d^3\mathbf{r}_A \phi_F^* \mathbf{r}_i \phi_I$$

by making use of the expansion

$$\frac{1}{|\mathbf{r} - \mathbf{r}_i|} \approx \frac{1}{r} + \frac{\mathbf{r} \cdot \mathbf{r}_i}{r^3}$$

Supplement 15-A

Lifetimes, Linewidths, and Resonances

In this supplement we do three things:

- (a) We discuss a somewhat improved treatment of transition rates that shows how the exponential decay behavior comes about, following the general approach of V. Weisskopf and E. P. Wigner.
- (b) We show how the Lorentzian shape for the linewidth comes about.
- (c) We show that the scattering amplitude for a photon by the atom in its ground state peaks strongly when the energy of the incident photon is equal to the (shifted) energy of the excited state.

To simplify the problem as much as possible, we consider an atom with just two levels, the ground state, with energy 0, and a single excited state, with energy E . The two states are coupled to the electromagnetic field, which we will take to be scalar, so that no polarization vectors appear. We will only consider the subset of eigenstates of H_0 consisting of the excited state ϕ_1 , for which

$$H_0|\phi_1\rangle = E|\phi_1\rangle \quad (15A-1)$$

and of the ground state + one photon, $\phi(\mathbf{k})$, for which

$$H_0|\phi(\mathbf{k})\rangle = \varepsilon(\mathbf{k})|\phi(\mathbf{k})\rangle \quad (15A-2)$$

and limit ourselves to these in an expansion of an arbitrary function. This is certainly justified when the coupling between the two states, ϕ_1 and $\phi(\mathbf{k})$ through the potential V , is small, as in electromagnetic coupling, since then the influence of two-, three-, . . . , photon states is negligible. Note that

$$\langle\phi_1|\phi(\mathbf{k})\rangle = 0 \quad (15A-3)$$

even when the $\mathbf{k}$ is such that the energies $\varepsilon(\mathbf{k})$ and E are the same. The states are orthogonal because one has a photon in it and the other does not, and because for one of them the atom is in an excited state, and for the other it is not.

The solution of the equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = (H_0 + V) |\psi(t)\rangle \quad (15A-4)$$

may be written in terms of the complete set

$$|\psi(t)\rangle = a(t)|\phi_1\rangle e^{-Et/\hbar} + \int d^3\mathbf{k} b(\mathbf{k}, t) |\phi(\mathbf{k})\rangle e^{-i\varepsilon(\mathbf{k})t/\hbar} \quad (15A-5)$$

When this is inserted into (15A-4),

$$\begin{aligned}
 & i\hbar \frac{da}{dt} e^{-iEt/\hbar} |\phi_1\rangle + Eae^{-iEt/\hbar} |\phi_1\rangle \\
 & + i\hbar \int d^3\mathbf{k} \frac{db(\mathbf{k}, t)}{dt} e^{-i\varepsilon(\mathbf{k})t/\hbar} |\phi(\mathbf{k})\rangle + \int d^3\mathbf{k} \varepsilon(\mathbf{k}) b(\mathbf{k}, t) e^{-i\varepsilon(\mathbf{k})t/\hbar} |\phi(\mathbf{k})\rangle \\
 & = Ea(t)e^{-iEt/\hbar} |\phi_1\rangle + \int d^3\mathbf{k} \varepsilon(\mathbf{k}) b(\mathbf{k}, t) e^{-i\varepsilon(\mathbf{k})t/\hbar} |\phi(\mathbf{k})\rangle \\
 & + a(t)e^{-iEt/\hbar} |\phi_1\rangle + \int d^3\mathbf{k} b(\mathbf{k}, t) e^{-i\varepsilon(\mathbf{k})t/\hbar} |\phi(\mathbf{k})\rangle
 \end{aligned}$$

results. If we take the scalar product with $\langle\phi_1|$, we get

$$i\hbar \frac{da}{dt} = a(t)\langle\phi_1|V|\phi_1\rangle + \int d^3\mathbf{k} b(\mathbf{k}, t) e^{-i[\varepsilon(\mathbf{k})-E]t/\hbar} \langle\phi_1|V|\phi(\mathbf{k})\rangle$$

Since V , acting on a state, is supposed to change the photon number by one, $\langle\phi_1|V|\phi_1\rangle = 0$. With the notation

$$\begin{aligned}
 \varepsilon(\mathbf{k}) - E &= \hbar\omega(\mathbf{k}) \\
 \langle\phi_1|V|\phi(\mathbf{k})\rangle &= M(\mathbf{k})
 \end{aligned} \tag{15A-6}$$

the equation becomes

$$i\hbar \frac{da(t)}{dt} = \int d^3\mathbf{k} b(\mathbf{k}, t) e^{-i\omega(\mathbf{k})t} M(\mathbf{k}) \tag{15A-7}$$

If we take the scalar product with $\langle\phi(\mathbf{q})|$, and again use photon counting to set $\langle\phi(\mathbf{q})|V|\phi(\mathbf{k})\rangle = 0$, we get, after a little manipulation, using a normalization that

$$\langle\phi(\mathbf{q})|\phi(\mathbf{k})\rangle = \delta(\mathbf{k} - \mathbf{q}) \tag{15A-8}$$

the equation

$$i\hbar \frac{db(\mathbf{q}, t)}{dt} = a(t) e^{i\omega(\mathbf{q})t} M^*(\mathbf{q}) \tag{15A-9}$$

Since $b(\mathbf{k}, 0) = 0$ if the excited state is occupied at $t = 0$, a solution of this equation is

$$b(\mathbf{k}, t) = \frac{1}{i\hbar} M^*(\mathbf{k}) \int_0^t dt' e^{i\omega(\mathbf{k})t'} a(t') \tag{15A-10}$$

$$\begin{aligned}
 \frac{da(t)}{dt} &= -\frac{1}{\hbar^2} \int d^3\mathbf{k} |M(\mathbf{k})|^2 e^{-i\omega(\mathbf{k})t} \int_0^t dt' a(t') e^{i\omega(\mathbf{k})t'} \\
 &= -\frac{1}{\hbar^2} \int d^3\mathbf{k} |M(\mathbf{k})|^2 \int_0^t dt' a(t') e^{-i\omega(\mathbf{k})(t-t')}
 \end{aligned} \tag{15A-11}$$

This is a complicated equation, and we will make some approximations to learn about the behavior of $a(t)$. We note that the oscillating exponential will average out to zero unless $t - t'$ is small. We therefore make the approximation that $a(t')$ can be replaced by $a(t)$. This turns a complicated integral equation into a simple ordinary differential equation.

$$\frac{1}{a(t)} \frac{da(t)}{dt} = -\frac{1}{\hbar^2} \int d^3\mathbf{k} |M(\mathbf{k})|^2 \int_0^t dt' e^{-i\omega(\mathbf{k})(t-t')} \tag{15A-12}$$

With the substitution $t - t' = s$ the integral over t' becomes $\int_0^1 ds e^{-i\omega(\mathbf{k})s}$. We now look at the behavior of $a(t)$ for large t . We use the fact that

$$\begin{aligned} \int_0^\infty ds e^{-i\omega(\mathbf{k})s} &= \lim_{\varepsilon \rightarrow 0^+} \int_0^\infty ds e^{-i(\omega(\mathbf{k}) - i\varepsilon)s} = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{i(\omega(\mathbf{k}) - i\varepsilon)} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{1}{i} \frac{\omega(\mathbf{k}) + i\varepsilon}{\omega^2(\mathbf{k}) + \varepsilon^2} \end{aligned} \quad (15A-13)$$

and the relation

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon}{\omega^2 + \varepsilon^2} = \pi \delta(\omega(\mathbf{k})) \quad (15A-14)$$

to obtain for the right hand side of (15A-12)

$$-\frac{\pi}{\hbar^2} \int d^3\mathbf{k} |M(\mathbf{k})|^2 \delta(\omega(\mathbf{k})) - \frac{i}{\hbar} \int d^3\mathbf{k} \frac{|\mathbf{M}(\mathbf{k})|^2}{\hbar\omega(\mathbf{k})} \quad (15A-15)$$

The solution of (15A-12) is therefore

$$a(t) = a(0) e^{-\Gamma/2 - i\Delta/\hbar} \quad (15A-16)$$

where

$$\Gamma = \frac{2\pi}{\hbar^2} \int d^3\mathbf{k} |\mathbf{M}(\mathbf{k})|^2 \delta(\omega(\mathbf{k})) = \frac{2\pi}{\hbar} \int d^3\mathbf{k} |\mathbf{M}(\mathbf{k})|^2 \delta(\hbar\omega(\mathbf{k})) \quad (15A-17)$$

and

$$\Delta = \int d^3\mathbf{k} \frac{|\mathbf{M}(\mathbf{k})|^2}{\hbar\omega(\mathbf{k})} \quad (15A-18)$$

The probability of finding the system in its initial state after a long time is

$$|\langle \phi_1 | \psi(t) \rangle|^2 = |a(t)|^2 = e^{-\Gamma t} \quad (15A-19)$$

This is the expected exponential decay form, with Γ coinciding with the second order perturbation calculation of the decay rate.

Another quantity of interest is the probability that the state $|\Psi(t)\rangle$ ends up in the state $|\phi(\mathbf{k})\rangle$ as t approaches infinity. This is given by $|b(\mathbf{k}, \infty)|^2$. This can be obtained from (15A-10). We can approximate this by using (15A-16). With the simplifying notation $z = \Gamma/2 + i\Delta/\hbar$, we get

$$\begin{aligned} b(\mathbf{k}, \infty) &= \frac{1}{i\hbar} M^*(\mathbf{k}) \int_0^\infty dt e^{(-z + i\omega(\mathbf{k}))t} = \frac{M^*(\mathbf{k})}{i\hbar} \frac{1}{z - i\omega(\mathbf{k})} \\ &= \frac{M^*(\mathbf{k})}{\hbar\omega(\mathbf{k}) - \Delta + i\hbar\Gamma/2} \end{aligned} \quad (15A-20)$$

The absolute square of this is

$$|b(\mathbf{k}, \infty)|^2 = \frac{|\mathbf{M}(\mathbf{k})|^2}{(\hbar\omega(\mathbf{k}) - \Delta)^2 + (\hbar\Gamma/2)^2} \quad (15A-21)$$

yields the Lorentzian shape for the linewidth; that is, the photon energy is centered about the (shifted) energy of the excited level, with the width described by $\hbar\gamma/2$. The energy shift is small, and usually ignored.

The same form appears in the scattering problem. Consider the scattering of a “photon” of momentum $\mathbf{k}_i$ by the atom in the ground state. The state of the system is again described by (15A-1), (15A-2), (15A-7), and (15A-9), except that initially, which here means at $t = -\infty$, the state is specifically given as $\phi(\mathbf{k}_i)$, so that

$$b(\mathbf{q}, t) = \delta(\mathbf{q} - \mathbf{k}_i) \quad \text{at } t = -\infty \quad (15A-22)$$

Hence the integration of (15A-9) gives

$$b(\mathbf{q}, t) = \delta(\mathbf{q} - \mathbf{k}_i) + \frac{1}{i\hbar} M^*(\mathbf{q}) \int_{-\infty}^t dt' a(t') e^{i\omega(\mathbf{q})t'} \quad (15A-23)$$

The quantity of interest is the amplitude for a transition into a final state in which the photon has momentum $\mathbf{k}_f$ at $t = +\infty$; that is, it is

$$\begin{aligned} \langle \phi(\mathbf{k}_f) | \psi(+\infty) \rangle &= b(\mathbf{k}_f, +\infty) \\ &= \delta(\mathbf{k}_f - \mathbf{k}_i) - \frac{i}{\hbar} M^*(\mathbf{k}_f) \int_{-\infty}^{\infty} dt' a(t') e^{i\omega_f t'} \end{aligned} \quad (15A-24)$$

$$(\omega_f \equiv \omega(\mathbf{k}_f))$$

using the previous equation.

The calculation of this quantity is rather tedious, and does not teach us any physics. The final result, though, is interesting. For scattering away from the forward direction, so that $\mathbf{k}_f \neq \mathbf{k}_i$, we find the result

$$b(\mathbf{k}_f, \infty) = \frac{-2\pi i M(\mathbf{k}_i) M^*(\mathbf{k}_f) \delta(\hbar\omega_i - \hbar\omega_f)}{\varepsilon(\mathbf{k}_i) - E + i\pi \int d^3\mathbf{k} |\mathbf{M}(\mathbf{k})|^2 \delta(\varepsilon(\mathbf{k}_i) - \varepsilon(\mathbf{k}))} \quad (15A-25)$$

The formula as it stands has been approximated by the neglect of a small real contribution to the denominator, whose effect is to shift the energy of the excited atom from E to $E + \Delta E$. What is of interest to us is that when the energy of the incident photon approaches that of the excited state of the atom E , the amplitude peaks very strongly. We have an example of *resonant scattering*. This is a quantitative justification of the remarks we made at the end of our discussion of *autoionization* in chapter 14.

Supplement 15-B

The Interaction Picture

For the discussion of systems involving only two or three levels, it is particularly convenient to use a description of the time evolution of the system that lies between the Schrödinger picture and the Heisenberg picture, both of which were discussed in Chapter 6. Let us start with the Schrödinger equation, which reads

$$\frac{d}{dt} |\psi(t)\rangle = -\frac{i}{\hbar} H |\psi(t)\rangle = -\frac{i}{\hbar} (H_0 + H_1) |\psi(t)\rangle \quad (15B-1)$$

We can write this in the form

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle \quad (15B-2)$$

where

$$\frac{d}{dt} U(t) = -\frac{i}{\hbar} (H_0 + H_1) U(t) \quad (15B-3)$$

The initial condition is $U(0) = 1$.

The procedure calls for the definition of a new state vector $|\psi_I(t)\rangle$ defined by

$$|\psi_I(t)\rangle = e^{iH_0 t/\hbar} |\psi(t)\rangle \quad (15B-4)$$

It follows that

$$\begin{aligned} \frac{d}{dt} |\psi_I(t)\rangle &= \frac{i}{\hbar} H_0 |\psi_I(t)\rangle + e^{iH_0 t/\hbar} \left(-\frac{i}{\hbar} \right) (H_0 + H_1) |\psi(t)\rangle \\ &= e^{iH_0 t/\hbar} \left(-\frac{i}{\hbar} \right) H_1 e^{-iH_0 t/\hbar} |\psi_I(t)\rangle \end{aligned}$$

If we now define

$$V(t) = e^{iH_0 t/\hbar} H_1 e^{-iH_0 t/\hbar} \quad (15B-5)$$

we end up with the equation

$$\frac{d}{dt} |\psi_I(t)\rangle = \left(-\frac{i}{\hbar} \right) V(t) |\psi_I(t)\rangle \quad (15B-6)$$

Solving this equation is not trivial, and in general the best one can do is to find a solution in terms of a power series in $V(t)$. The formal procedure for solving this in a way that incorporates the initial condition

$$|\psi_I(0)\rangle = |\phi\rangle \quad (15B-7)$$

is to write

$$|\psi_I(t)\rangle = U_I(t) |\phi\rangle \quad (15B-8)$$

Equation (15B-6) takes the form

$$\frac{dU_I(t)}{dt} = -\frac{i}{\hbar} V(t)U_I(t) \quad (15B-9)$$

Since $U_I(0) = 1$, we can convert the differential equation into an integral equation

$$U_I(t) = 1 - \frac{i}{\hbar} \int_0^t dt' V(t')U_I(t') \quad (15B-10)$$

This can be solved by iteration. In the first step we replace the U_I under the integral by 1. In the second step, we take the $U_I(t)$ so obtained and insert it on the right-hand side, and so on. We thus get

$$U_I(t) = 1 - \frac{i}{\hbar} \int_0^t dt' V(t') + \left(-\frac{i}{\hbar}\right)^2 \int_0^t dt' V(t') \int_0^{t'} dt'' V(t'') + \cdots \quad (15B-11)$$

This is a nice compact form, but working out the integrals in the second term is still very tedious. That is all we have to say about this, other than to say that the first-order expression is very handy for dealing with two- and three-level systems, as we shall see in Chapter 18.

Chapter 16

The Interaction of Charged Particles with the Electromagnetic Field

The most important application of quantum mechanics dealt with in this book involves the interaction between electrons (or other charged particles) and the electromagnetic field. In this chapter we deal with the formulation of the Schrödinger equation in the presence of a general electromagnetic field, and then study the special case of electrons in a constant, uniform magnetic field.

16-1 CLASSICAL ELECTRODYNAMICS

Maxwell's equations read

$$\begin{aligned}\nabla \cdot \mathbf{B}(\mathbf{r}, t) &= 0 \\ \nabla \cdot \mathbf{E}(\mathbf{r}, t) &= \frac{1}{\epsilon_0} \rho(\mathbf{r}, t) \\ \nabla \times \mathbf{B}(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} &= \mu_0 \mathbf{j}(\mathbf{r}, t) \\ \nabla \times \mathbf{E}(\mathbf{r}, t) + \frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} &= 0\end{aligned}\tag{16-1}$$

These equations lead to

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0\tag{16-2}$$

which expresses charge conservation. If we integrate (16-2) over an arbitrary volume V , then this reads

$$\frac{d}{dt} \int_V d^3r \rho(\mathbf{r}, t) = - \int_S dS \mathbf{n} \cdot \mathbf{j}(\mathbf{r}, t)$$

which states that the rate of change of charge in the volume V is equal to the net current flowing out through the surface S that bounds the volume V .

A point electron of mass m_e and charge $-e$ obeys the Lorentz force equation

$$m_e \frac{d^2 \mathbf{r}}{dt^2} = -e[\mathbf{E}(\mathbf{r}, t) + \mathbf{v} \times \mathbf{B}(\mathbf{r}, t)]\tag{16-3}$$

The transition to quantum mechanics requires the construction of a Hamiltonian. To do this, it is necessary to introduce the *potentials* for the electromagnetic system. We introduce $\mathbf{A}(\mathbf{r}, t)$ and $\phi(\mathbf{r}, t)$ as follows

$$\begin{aligned}\mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{A}(\mathbf{r}, t) \\ \mathbf{E}(\mathbf{r}, t) &= -\frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} - \nabla \phi(\mathbf{r}, t)\end{aligned}\quad (16-4)$$

so that the first and last of the equations (16-1) are automatically satisfied. The potentials are not uniquely defined. The first of the above equations is unchanged if we make a transformation to a new *vector potential*

$$\mathbf{A}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t) - \nabla g(\mathbf{r}, t) \quad (16-5)$$

since $\nabla \times \nabla g(\mathbf{r}, t) = 0$. The electric field will be unchanged if, in addition to making the change in (16-5), we change ϕ to ϕ' according to

$$\phi'(\mathbf{r}, t) = \phi(\mathbf{r}, t) + \frac{\partial g(\mathbf{r}, t)}{\partial t} \quad (16-6)$$

This invariance, known as *invariance under gauge transformations*, gives us the freedom to define the potentials in various ways, to suit our convenience. The Maxwell equations now read

$$-\nabla^2 \phi(\mathbf{r}, t) - \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}(\mathbf{r}, t)) = \frac{1}{\epsilon_0} \rho(\mathbf{r}, t) \quad (16-7)$$

and

$$\nabla \times (\nabla \times \mathbf{A}(\mathbf{r}, t)) + \frac{1}{c^2} \frac{\partial^2 \mathbf{A}(\mathbf{r}, t)}{\partial t^2} + \frac{1}{c^2} \nabla \frac{\partial \phi(\mathbf{r}, t)}{\partial t} = \mu_0 \mathbf{j}(\mathbf{r}, t) \quad (16-8)$$

where we have used $\epsilon_0 \mu_0 = 1/c^2$. This may be rewritten as

$$-\nabla^2 \mathbf{A}(\mathbf{r}, t) + \frac{1}{c^2} \frac{\partial^2 \mathbf{A}(\mathbf{r}, t)}{\partial t^2} + \nabla \left(\nabla \cdot \mathbf{A}(\mathbf{r}, t) + \frac{1}{c^2} \frac{\partial \phi(\mathbf{r}, t)}{\partial t} \right) = \mu_0 \mathbf{j}(\mathbf{r}, t) \quad (16-9)$$

If the charge distribution is static—that is, $\rho(\mathbf{r})$ is independent of time—it is convenient to choose the function $g(\mathbf{r}, t)$ such that

$$\nabla \cdot \mathbf{A}(\mathbf{r}, t) = 0 \quad (16-10)$$

This choice is given the name of *Coulomb gauge*. In this case we have

$$\nabla^2 \phi(\mathbf{r}) = \frac{1}{\epsilon_0} \rho(\mathbf{r}) \quad (16-11)$$

a time-independent scalar potential. The equation for the vector potential in this gauge reads

$$-\nabla^2 \mathbf{A}(\mathbf{r}, t) + \frac{1}{c^2} \frac{\partial^2 \mathbf{A}(\mathbf{r}, t)}{\partial t^2} = \mu_0 \mathbf{j}(\mathbf{r}, t) \quad (16-12)$$

When the charge distribution is not static, it is more convenient to choose the so-called *Lorentz gauge*, for which

$$\nabla \cdot \mathbf{A}(\mathbf{r}, t) + \frac{1}{c^2} \frac{\partial \phi(\mathbf{r}, t)}{\partial t} = 0 \quad (16-13)$$

This leaves the equation for the vector potential (16-12) unaltered. The scalar potential now also obeys a wave equation. A technical point worth noting is that the vector identity,

$$\nabla \times (\nabla \times \mathbf{A}) = -\nabla^2 \mathbf{A} + \nabla(\nabla \cdot \mathbf{A})$$

used to obtain (16-9) is only valid in cartesian coordinates. This means that $\nabla^2 \mathbf{A}(\mathbf{r}, t)$ as it appears, must be calculated in terms of x , y , and z .

The transition to quantum mechanics requires a Hamiltonian formulation of the equation of motion. In the absence of interaction with the electromagnetic field, it is easily seen that the Hamiltonian equations

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial x_i} \end{aligned} \quad (16-14)$$

with

$$H = \frac{p^2}{2m_e} + V(r) \quad (16-15)$$

yields

$$m_e \frac{d^2 x_i}{dt^2} = -\frac{\partial V(r)}{\partial x_i} \quad (16-16)$$

The Hamiltonian for the interaction of an electron with an external electromagnetic field, represented by the potentials $(\mathbf{A}(\mathbf{r}, t), \phi(\mathbf{r}, t))$, is taken to be

$$H = \frac{1}{2m_e} (\mathbf{p} + e\mathbf{A}(\mathbf{r}, t))^2 - e\phi(\mathbf{r}, t) \quad (16-17)$$

The Hamiltonian equations of motion are

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} = \frac{1}{m_e} (p_i + eA_i(\mathbf{r}, t)) \quad (16-18)$$

and

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i} = -\frac{1}{m_e} (p_k + eA_k) \frac{\partial A_k}{\partial x_i} + e \frac{\partial \phi}{\partial x_i} \quad (16-19)$$

From this it follows that

$$\begin{aligned} m_e \frac{d^2 x_i}{dt^2} &= \frac{d}{dt} (p_i + eA_i(\mathbf{r}, t)) = \frac{dp_i}{dt} + e \left(\frac{\partial A_i}{\partial t} + \frac{\partial A_i}{\partial x_k} \frac{dx_k}{dt} \right) \\ &= -\frac{e}{m_e} \left(m_e \frac{dx_k}{dt} \right) \frac{\partial A_k}{\partial x_i} + e \frac{\partial \phi}{\partial x_i} + e \frac{\partial A_i}{\partial t} + e \frac{\partial A_i}{\partial x_k} \frac{dx_k}{dt} \\ &= e \left(\frac{\partial \phi}{\partial x_i} + \frac{\partial A_i}{\partial t} \right) + e \frac{dx_k}{dt} \left(\frac{\partial A_i}{\partial x_k} - \frac{\partial A_k}{\partial x_i} \right) \\ &= -eE_i + e(\mathbf{v} \times (-\nabla \times \mathbf{A}))_i \\ &= -eE_i - e(\mathbf{v} \times \mathbf{B})_i \end{aligned} \quad (16-20)$$

This confirms that the Hamiltonian in (16-7) correctly describes the classical motion.

16-2 THE SCHRÖDINGER EQUATION FOR AN ELECTRON IN INTERACTION WITH AN ELECTROMAGNETIC FIELD

The Schrödinger equation takes the form

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = H\psi = \left[\frac{1}{2m_e} (-i\hbar \nabla + e\mathbf{A}(\mathbf{r}, t))^2 - e\phi(\mathbf{r}, t) \right] \psi(\mathbf{r}, t) \quad (16-21)$$

where we have replaced the operator $\mathbf{p}$ by $-i\hbar \nabla$. Let us first check what happens to this equation when we make a gauge transformation on the potentials, such as (16-5) and (16-6). In terms of the primed potentials the equation becomes

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = \left[\frac{1}{2m_e} (-i\hbar \nabla + e\mathbf{A}'(\mathbf{r}, t) + e\nabla g(\mathbf{r}, t))^2 - e\phi'(\mathbf{r}, t) - e \frac{\partial g(\mathbf{r}, t)}{\partial t} \right] \psi(\mathbf{r}, t)$$

This looks like a different equation. We now show that by making an appropriate change from ψ to a newly defined ψ' we can restore the equation to its original form, and thus show the *gauge invariance* of the Schrödinger equation. Since we want to preserve the absolute square of the wave function, we define

$$\psi'(\mathbf{r}, t) = e^{i\Lambda(\mathbf{r}, t)} \psi(\mathbf{r}, t) \quad (16-22)$$

We see that

$$\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial t} (e^{-i\Lambda} \psi') = -i \frac{\partial \Lambda}{\partial t} \psi + e^{-i\Lambda} \frac{\partial \psi'}{\partial t}$$

and

$$-i\hbar \nabla \psi = -i\hbar \nabla (e^{-i\Lambda} \psi') = -\hbar (\nabla \Lambda) \psi - e^{-i\Lambda} \nabla \psi'$$

We see from this that we recover the original Schrödinger equation in terms of $\mathbf{A}'$, ϕ' , ψ' , provided we choose

$$\Lambda(\mathbf{r}, t) = \frac{e}{\hbar} g(\mathbf{r}, t) \quad (16-23)$$

To continue, we assume that the scalar potential is independent of time. We work in the *Coulomb gauge* specified by

$$\nabla \cdot \mathbf{A}(\mathbf{r}, t) = 0 \quad (16-24)$$

Thus (16-21), which has the form

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m_e} \nabla^2 \psi - \frac{ie\hbar}{m_e} \mathbf{A} \cdot \nabla \psi - \frac{ie\hbar}{m_e} (\nabla \cdot \mathbf{A}) \psi + \frac{e^2}{2m_e} \mathbf{A}^2 \psi - e\phi \psi$$

becomes

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m_e} \nabla^2 \psi - \frac{ie\hbar}{m_e} \mathbf{A} \cdot \nabla \psi + \frac{e^2}{2m_e} \mathbf{A}^2 \psi - e\phi \psi \quad (16-25)$$

Note that for a Coulomb potential,

$$\phi(\mathbf{r}) = -\frac{Ze}{4\pi\epsilon_0 r} \quad (16-26)$$

In what follows we discuss the interaction of an electron with a constant magnetic field. This is a rather rich subject, with many important physical applications.

Since the magnetic field is constant, we may write

$$\psi(\mathbf{r}, t) = e^{-iEt\hbar}\psi(\mathbf{r}) \quad (16-27)$$

The Schrödinger equation becomes

$$-\frac{\hbar^2}{2m_e}\nabla^2\psi - \frac{ie\hbar}{m_e}\mathbf{A}\cdot\nabla\psi + \frac{e^2}{2m_e}\mathbf{A}^2\psi - e\phi\psi = E\psi \quad (16-28)$$

16-3 THE CONSTANT MAGNETIC FIELD

For a constant, uniform magnetic field $\mathbf{B}$, we may take¹

$$\mathbf{A} = -\frac{1}{2}\mathbf{r} \times \mathbf{B} \quad (16-29)$$

It is simple to check that $\nabla \times \mathbf{A} = \mathbf{B}$. Thus, in the Schrödinger equation above, the second term becomes

$$-\frac{ie\hbar}{m_e}\left(-\frac{1}{2}\mathbf{r} \times \mathbf{B}\right)\cdot\nabla\psi = -\frac{ie\hbar}{2m_e}\mathbf{B}\cdot\mathbf{r} \times \nabla\psi = \frac{e}{2m_e}\mathbf{B}\cdot\mathbf{L}\psi \quad (16-30)$$

This is a consequence of the relation

$$-i\hbar\mathbf{r} \times \nabla\psi = \mathbf{r} \times \mathbf{p}_{\text{op}}\psi = \mathbf{L}\psi$$

The term quadratic in $\mathbf{A}$ takes the form

$$\frac{e^2}{2m_e}\left(-\frac{1}{2}\mathbf{r} \times \mathbf{B}\right)^2\psi = \frac{e^2}{8m_e}(r^2B^2 - (\mathbf{r}\cdot\mathbf{B})^2)\psi = \frac{e^2B^2}{8m_e}(x^2 + y^2)\psi \quad (16-31)$$

if the z-axis is defined by the direction of $\mathbf{B}$, so that $\mathbf{B} = (0, 0, B)$. This term looks like a two-dimensional harmonic oscillator potential.

Let us compare the magnitudes of the two terms. In making our estimate we choose $\langle L_z \rangle$ of order $\hbar$, and $(x^2 + y^2)$ to be of order a_0^2 , where a_0 is the Bohr radius. This means that

$$\begin{aligned} \frac{e^2B^2a_0^2/8m_e}{eB\hbar/2m_e} &= \frac{eBa_0^2}{4\hbar} = \frac{(1.6 \times 10^{-19} \text{ C})(0.5 \times 10^{-10} \text{ m})^2}{4(1.05 \times 10^{-34} \text{ J}\cdot\text{s})} B(\text{tesla}) \\ &= \frac{B(\text{tesla})}{10^6} \end{aligned}$$

For B of the order of 1 tesla, the quadratic term is negligible. The term linear in B is also small compared with typical atomic splittings. Comparison with the Coulomb energy yields

$$\frac{eB\hbar/2m_e}{m_e c^2 \alpha^2 / 2} = \frac{e\hbar}{(m_e c \alpha)^2} B = \frac{B(\text{tesla})}{2.3 \times 10^5}$$

The quadratic term can become important under certain conditions. One of them is when the magnetic field is very strong. Fields as large as 10^8 T may exist on the surface of neutron stars, and such fields overwhelm the electric fields due to the electrostatic attraction between electrons and nuclei. In such an environment the structure of atoms is radically

¹This choice is not unique, since we may add the gradient of any function Λ without changing $\mathbf{B}$. This choice is convenient for our purposes.

altered. The other condition is when the $(x^2 + y^2)$ term in the interaction is of macroscopic size. This happens for motion of an electron in a synchrotron. Here both the radius of the orbit and L are huge on an atomic scale.

If we wanted to take spin into account, we would only have to alter the term linear in $\mathbf{B}$ in the Schrödinger equation so that $\mathbf{L} \cdot \mathbf{B}$ is replaced by $(2\mathbf{S} + \mathbf{L}) \cdot \mathbf{B}$. In what follows, we ignore the spin.

The presence of $(x^2 + y^2)$ in the potential suggest the use of cylindrical coordinates. With

$$\begin{aligned} x &= \rho \cos \phi \\ y &= \rho \sin \phi \end{aligned} \quad (16-32)$$

we follow the procedure outlined at the beginning of Supplement 7-B [www.wiley.com/college/gasiorowicz] to arrive at

$$\begin{aligned} \frac{\partial}{\partial x} &= \cos \phi \frac{\partial}{\partial \rho} - \frac{\sin \phi}{\rho} \frac{\partial}{\partial \phi} \\ \frac{\partial}{\partial y} &= \sin \phi \frac{\partial}{\partial \rho} + \frac{\cos \phi}{\rho} \frac{\partial}{\partial \phi} \end{aligned} \quad (16-33)$$

and hence

$$\nabla^2 = \frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \quad (16-34)$$

If we now write

$$\psi(\mathbf{r}) = u_m(\rho) e^{im\phi} e^{ikz} \quad (16-35)$$

We find that the differential equation satisfied by $u_m(\rho)$ is

$$\frac{d^2 u_m}{d\rho^2} + \frac{1}{\rho} \frac{du_m}{d\rho} - \frac{m^2}{\rho^2} u_m - \frac{e^2 B^2}{4\hbar^2} \rho^2 u_m + \left(\frac{2m_e E}{\hbar^2} - \frac{eB\hbar m}{\hbar^2} - k^2 \right) u_m = 0 \quad (16-36)$$

If we introduce the variable

$$x = \sqrt{\frac{eB}{2\hbar}} \rho \quad (16-37)$$

we can rewrite the equation in the form

$$\frac{d^2 u_m}{dx^2} + \frac{1}{x} \frac{du_m}{dx} - \frac{m^2}{x^2} u_m - x^2 u_m + \lambda u_m = 0 \quad (16-38)$$

where

$$\lambda = \frac{4m_e}{eB\hbar} \left(E - \frac{\hbar^2 k^2}{2m_e} \right) - 2m \quad (16-39)$$

It is fairly straightforward to determine that (a) the behavior of $u(x)$ at infinity, determined from

$$\frac{d^2 u}{dx^2} - x^2 u \approx 0$$

is $u(x) \sim e^{-x^2/2}$, and (b) the behavior of $u(x)$ near $x = 0$, determined from

$$\frac{d^2 u}{dx^2} + \frac{1}{x} \frac{du}{dx} - \frac{m^2}{x^2} u \approx 0$$

is $u(x) \sim x^{|m|}$. We thus write

$$u(x) = x^{|m|} e^{-x^2/2} G(x) \quad (16-40)$$

and determine the differential equation obeyed by $G(x)$. A little algebra leads to

$$\frac{d^2 G}{dx^2} + \left(\frac{2|m|+1}{x} - 2x \right) \frac{dG}{dx} + (\lambda - 2 - 2|m|) G = 0 \quad (16-41)$$

This can be brought into the same form as (8-27) if we change variables to

$$y = x^2 \quad (16-42)$$

The equation then takes the form

$$\frac{d^2 G}{dy^2} + \left(\frac{|m|+1}{y} - 1 \right) \frac{dG}{dy} + \frac{\lambda - 2 - 2|m|}{4y} G = 0 \quad (16-43)$$

We can now proceed as in Chapter 8. Comparison with (8-32) shows that we must have

$$\frac{1}{4} \lambda - \frac{1+|m|}{2} = n_r \quad (16-44)$$

as an eigenvalue condition, with $n_r = 0, 1, 2, 3, \dots$. This implies that $E - \hbar^2 k^2 / 2m_e$, the energy with the kinetic energy of free motion in the z -direction subtracted out, is given by

$$E - \frac{\hbar^2 k^2}{2m_e} = \frac{eB\hbar}{2m_e} (2n_r + 1 + |m| + m) \quad (16-45)$$

and

$$G(y) = L_{n_r}^{|m|}(y) \quad (16-46)$$

Classical Limit

We limit our discussion to the classical limit. We first review classical theory. Recall that we found the velocity in the form

$$\mathbf{v} = \frac{\mathbf{p} + e\mathbf{A}}{m_e} \quad (16-47)$$

With $\mathbf{A} = -\frac{1}{2}\mathbf{r} \times \mathbf{B}$ we find

$$\begin{aligned} m_e \mathbf{r} \times \mathbf{v} &= \mathbf{r} \times \mathbf{p} + e\mathbf{r} \times \left(-\frac{1}{2} \mathbf{r} \times \mathbf{B} \right) \\ &= \mathbf{L} - \frac{e}{2} [\mathbf{r}(\mathbf{r} \cdot \mathbf{B}) - r^2 \mathbf{B}] \end{aligned} \quad (16-48)$$

We take the z -component of the equation to obtain

$$m_e (\mathbf{r} \times \mathbf{v})_z = L_z + \frac{eB}{2} (x^2 + y^2)$$

that is,

$$m_e \rho v = L_z + \frac{eB}{2} \rho^2 \quad (16-49)$$

The force on an electron is given by $\mathbf{F} = -e\mathbf{v} \times \mathbf{B}$, so that for circular motion we have

$$m_e \frac{v^2}{\rho} = evB \quad (16-50)$$

This, with (16-49) yields, after a little algebra,

$$\frac{1}{2} m_e v^2 = \frac{eB}{m_e} L_z \quad (16-51)$$

and

$$\rho = \sqrt{\frac{2L_z}{eB}} \quad (16-52)$$

We now return to the expression for the energy, (16-45). Because of the smallness of $\hbar$, the energy can only be of macroscopic size for reasonable B , if $(2n_r + 1 + |m| + m)$ is very large. We have two cases: (a) If $m < 0$, this implies that n_r is very large. Now n_r determines the degree of the polynomial $L_{n_r}^{|m|}(y)$ —that is, the number of the zeros in the function²—and if that is very large, the function cannot be large for some small range of y where the classical orbit would be located. (b) If $m > 0$, the coefficient is $(2n_r + 1 + 2m)$, and this can be large, with n_r small, provided that m is large. The energy now is

$$E - \frac{\hbar^2 k^2}{2m_e} \cong \left(\frac{eB}{m_e} \right) \hbar m \quad (16-53)$$

in agreement with the classical result. Note that

$$L_z = \hbar m \quad (16-54)$$

is positive, as expected.

We can also show that the radius of the orbit, as determined by the peaking of the radial probability distribution, corresponds to the classical value. Let us take $n_r = 0$. In that case $L_{n_r}^{|m|}(y)$ is just a constant, and the square of the wave function is according to (16-40),

$$P(x) = x^{2|m|} e^{-x^2} \quad (16-55)$$

This has a maximum where

$$\frac{dP}{dx} = (2|m|x^{2|m|-1} - 2x^{2|m|+1}) e^{-x^2} = 0$$

that is, at

$$x = \sqrt{|m|} \quad (16-56)$$

which yields

$$\rho = \left(\frac{2}{eB} \hbar m \right)^{1/2} \quad (16-57)$$

This problem provides a beautiful illustration of the correspondence principle.

16-4 LANDAU LEVELS

The choice $\mathbf{A} = (-yB/2, xB/2, 0)$ is not unique. The choice

$$\mathbf{A} = (0, Bx, 0) \quad (16-58)$$

²See the discussion at the beginning of the section on the Degeneracy of the Spectrum in Chapter 8.

leads to the same magnetic field. It differs by a simple gauge transformation from $\mathbf{A} = -\mathbf{r} \times \mathbf{B}/2$:

$$\left(\frac{-yB}{2}, \frac{xB}{2}, 0\right) = (0, Bx, 0) - \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)\left(\frac{yxB}{2}\right) \quad (16-59)$$

With this choice of vector potential, the Hamiltonian operator for an electron in a constant magnetic field takes the form

$$\begin{aligned} \frac{1}{2m_e}(\mathbf{p} + e\mathbf{A})^2 &= \frac{1}{2m_e}(p_x^2 + (p_y + eBx)^2 + p_z^2) \\ &= \frac{1}{2m_e}(p_x^2 + p_y^2 + p_z^2 + 2eBxp_y + e^2B^2x^2) \end{aligned} \quad (16-60)$$

It is clear that $[H, p_y] = 0$ and $[H, p_z] = 0$, so that we can construct functions that are simultaneous eigenfunctions of p_y, p_z as well as H . For simplicity we choose the state to be an eigenfunction of p_z with eigenvalue zero—that is, no motion outside of the x - y plane. If we write the eigenvalue of p_y as $\hbar k$, then the simultaneous eigenfunction takes the form

$$\psi(x, y) = e^{iky} v(x) \quad (16-61)$$

where $v(x)$ is a solution of the equation

$$\frac{1}{2m_e} \left(-\hbar^2 \frac{d^2}{dx^2} + e^2 B^2 \left(x + \frac{\hbar k}{eB} \right)^2 \right) v(x) = E v(x) \quad (16-62)$$

This is just the equation for a harmonic oscillator whose equilibrium point is shifted from $x = 0$ to $-x_0$ where $x_0 = \hbar k/eB$. We may thus write the solution as

$$\psi(x, y) = e^{ieBx_0 y/\hbar} u(x - x_0) \quad (16-63)$$

where $u(x)$ is the eigensolution of the harmonic oscillator centered at $x = 0$. Note that a comparison with the usual potential $\frac{1}{2}m_e\omega^2x^2$ shows that the frequency is

$$\omega = \frac{eB}{m_e} \quad (16-64)$$

and the energy eigenvalues are

$$E = \hbar\omega \left(n + \frac{1}{2} \right) \quad n = 0, 1, 2, \dots \quad (16-65)$$

The energy levels labeled by n are called Landau levels.

If the electron is confined to a strip whose x -dimension is L_1 and y -dimension is L_2 , then the boundary condition in the y -direction,

$$\psi(y) = \psi(y + L_2) \quad (16-66)$$

implies that

$$\frac{eBx_0}{\hbar} L_2 = 2\pi n^* \quad n^* = 0, 1, 2, \dots \quad (16-67)$$

Since

$$0 \leq x_0 \leq L_1 \quad (16-68)$$

we deduce that

$$0 \leq n^* \leq \frac{eB}{2\pi\hbar} L_1 L_2 \quad (16-69)$$

We can easily check that $\hbar/eB$ has the dimensions of an area. (A quick way is to see that v_eB appearing in the Lorentz force has the dimensions $[\text{MLT}^{-2}]$. Therefore eB has the dimensions $[\text{MT}^{-1}]$. With $\hbar$ having dimensions of $\text{J} \cdot \text{s}$ —that is, $[\text{ML}^2\text{T}^{-1}]$, we get our result.) If we define a *magnetic length* by

$$l_B^2 = \frac{\hbar}{eB} \quad (16-70)$$

(whose value is $l_B = \frac{256 \times 10^{-10}}{\sqrt{B(\text{tesla})}} \text{ m}$),

we can write

$$n_{\text{max}}^* = \frac{L_1 L_2}{2\pi l_B^2} \quad (16-71)$$

Since the area of the sample is L_1, L_2 , the number of states per unit area of a full Landau level is therefore

$$n_B = \frac{1}{2\pi l_B^2} \quad (16-72)$$

The energy spacing of the Landau levels is

$$\hbar\omega = \frac{eB\hbar}{m_e} = \frac{\hbar^2}{m_e l_B^2} \quad (16-73)$$

The occupation of Landau levels is of great importance in the theoretical understanding of the *Quantum Hall Effect*. We present a very much simplified discussion of this topic.

16-5 THE INTEGRAL QUANTUM HALL EFFECT

Consider a two-dimensional sample as shown in Fig. 16-1. The $\mathbf{B}$ field points in the z -direction and an electric field $\mathbf{E}$ is applied in the y -direction. Electrons will flow in the negative y -direction, with a current density proportional to the electric field and the conductivity, so that

$$j_y = \sigma_0 E_y \quad (16-74)$$

The conductivity is given by $\sigma_0 = n_e e^2 \tau_0 / m_e^*$. Here n_e is the electron density, m_e^* is the effective mass of the electron in the material, and τ_0 is a quantity with the dimensions of time, which can be interpreted as the average time between collisions of the electron with imperfections in the lattice.³ In the presence of the magnetic field the electrons experience a Lorentz force $\mathbf{F} = -e\mathbf{v} \times \mathbf{B}$. Since the current is related to the velocity by $\mathbf{j} = -n_e e\mathbf{v}$, the extra force due to the magnetic field is

$$\mathbf{F} = \frac{\mathbf{j} \times \mathbf{B}}{n_e}$$

This is equivalent to an additional electric field $\mathbf{E}' = -\mathbf{F}/e$, so that the current density is changed to

$$\mathbf{j} = \sigma_0 \left(\mathbf{E} - \frac{\mathbf{j} \times \mathbf{B}}{n_e e} \right) \quad (16-75)$$

³For a brief discussion of conductivity, see Bernstein, Fishbane, and Gasiorowicz, *loc. cit.*

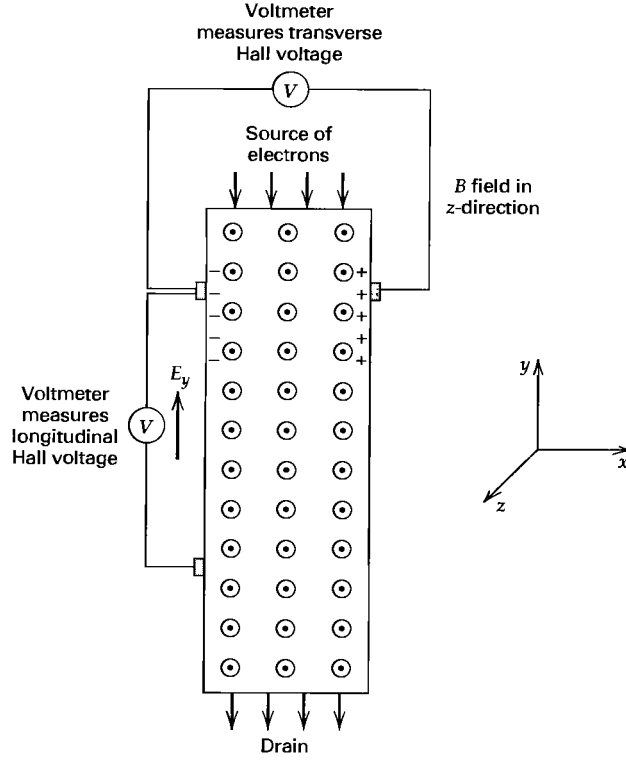


Figure 16-1 Schematics for measuring Hall voltage.

The components of this are

$$j_x = -\frac{\sigma_0 j_y B}{n_e e} = -\frac{e\tau_0}{m_e^*} j_y B \quad (16-76)$$

and

$$\begin{aligned} j_y &= \sigma_0 E_y + \frac{\sigma_0 j_x B}{n_e e} = \sigma_0 E_y + \frac{e\tau_0}{m_e^*} j_x B \\ &= \sigma_0 E_y - \left(\frac{e\tau_0}{m_e^*} \right)^2 j_y B^2 \end{aligned} \quad (16-77)$$

This leads to the following expressions for the current:

$$j_y = \frac{\sigma_0 E_y}{1 + (e\tau_0 B/m_e^*)^2} \quad (16-78)$$

and, after some manipulations,

$$j_x = \frac{n_e e E_y}{B} \left(1 - \frac{1}{1 + (e\tau_0 B/m_e^*)^2} \right) \quad (16-79)$$

The density of electron—that is, the number of electrons per unit area, n_e —may be written in terms of the maximum number of electrons that can be accommodated in a Landau level

$$n_e = f n_B = f \frac{eB}{2\pi\hbar} \quad (16-80)$$

This lead to

$$\begin{aligned}\frac{j_y}{E_y} &= \frac{\sigma}{1 + (e\tau_0 B/m_e^*)^2} \\ \frac{j_x}{E_y} &= f \frac{e^2}{2\pi\hbar} \left(1 - \frac{1}{1 + (e\tau_0 B/m_e^*)^2} \right)\end{aligned}\quad (16-81)$$

The density of electrons and the magnetic field B are under the control of the experimentalist. If n_e is fixed and B is varied, then the ratios in (16-81) can be measured as a function of B . Equivalently, if B is fixed and n_e varied, the ratios can be measured as functions of n_e . It was found by von Klitzing, Dorda, and Pepper in 1980 that the values of B such that f takes on integer values $f = 1, 2, 3, \dots$ lead to (a) vanishing values of j_y/E_y , and (b) values of $|j_x/E_y| = f(e^2/2\pi\hbar)$. A simple-minded explanation of the effect is that, when the Landau levels are filled, an electron cannot undergo elastic scattering, since it cannot recoil into another state of the same energy. It cannot be thermally excited into a higher (empty) Landau level, since at low temperatures (0.1 K) and large magnetic fields ($B \approx 10$ T), $kT \ll eB\hbar/m_e^*$. Effectively $\tau_0 \rightarrow \infty$, and the observed result follows from (16-81).

This is an oversimplified discussion, since it does not take into account a very large number of effects that occur in the non-ideal case. For example, some electrons are trapped at imperfections in the crystal lattice, Landau levels are not sharp because of thermal and impurity effects, and electron-electron interactions have been completely neglected. Nevertheless, when all of these complications are taken into account, it is still true that at critical values of B , j_x/E_y is an integral multiple of $e^2/2\pi\hbar$, to an accuracy of better than one part in 10 million.⁴

16-6 AN ADDITIONAL COMMENT ON GAUGE INVARIANCE

The relations between the electric and magnetic fields and the vector and scalar potentials appear to suggest that when the fields vanish, so do the potentials. This is not as simple a statement as appears on the face of it. Let us consider a constant magnetic field pointing in the z -direction

$$\mathbf{B} = k\mathbf{B}$$

confined to a cylindrical region $0 \leq \rho \leq a$. We may use Stokes' theorem

$$\oint d\ell \cdot \mathbf{A} = 2\pi\rho A_\phi = \int_{\text{surface}} d\mathbf{S} \cdot \mathbf{B} = \pi\rho^2 B$$

to see that $\mathbf{A}$ in cylindrical coordinates only has a component in the ϕ -direction, and that for $\rho \leq a$,

$$A_\phi = \frac{1}{2} B\rho \quad (16-82)$$

If we take our circular path about the field outside the radius a , we get

$$2\pi\rho A_\phi = \pi a^2 B$$

⁴The precise *quantization* actually follows from gauge invariance, as has been pointed out by R. Laughlin and B. Halperin. A presentation of this argument may be found in C. Kittel, *Introduction to Solid State Physics*, 6th Ed., John Wiley, New York, 1986.

so that for $\rho \geq a$,

$$A_\phi = \frac{B\pi a^2}{2\pi\rho} = \frac{\Phi}{2\pi\rho} \quad (16-83)$$

where Φ is the uniform magnetic flux contained in the cylinder $\rho \leq a$. We see that in the outer region the vector potential *does not vanish* even though the magnetic field does. In the region in which the magnetic field vanishes the vector potential may be written as the gradient of a function $g(\phi)$,

$$A_\phi = \frac{1}{\rho} \frac{dg(\phi)}{d\phi} \quad (16-84)$$

When we set this equal to $\frac{\Phi}{2\pi\rho}$, we get

$$g = \left(\frac{\Phi}{2\pi} \right) \phi \quad (16-85)$$

where ϕ is the polar angle. In classical electromagnetic theory we can make a gauge transformation that sets $A_\phi = 0$ —we just add to $\mathbf{A}$ the gradient of $-g$. In quantum mechanics this has some additional consequences. As noted in the material leading up to eq. (16-19), the Schrödinger wave function for an electron acquires an additional phase factor, which is

$$e^{i\Lambda} = e^{ieg\hbar} = e^{i(e\Phi/2\pi\hbar)\phi} \quad (16-86)$$

The phase factor has no consequences for $|\psi(\mathbf{r}, t)|^2$ but there are circumstances in which it is measurable. This was first pointed out by Y. Aharonov and D. Bohm, and is discussed in some detail in Supplement 16-A [www.wiley.com/college/gasiorowicz].

PROBLEMS

1. A particle of mass m in a three-dimensional harmonic oscillator of potential energy $m\omega^2 r^2/2$ has a spectrum given by

$$E = \hbar\omega(2n_r + l + 3/2)$$

where n_r is a radial quantum number ($n_r = 0, 1, 2, 3, \dots$) and l is the orbital angular momentum ($l = 0, 1, 2, 3, \dots$). Suppose the particle has charge q and the harmonic oscillator is placed in a weak magnetic field $\mathbf{B}$. Sketch the spectrum for the three lowest energy states.

2. Consider a particle of mass M attached to a rigid massless rod of fixed length R whose other end is fixed at the origin. The rod is free to rotate about its fixed point.

(a) Give an argument why the Hamiltonian for the system may be written as

$$H = \frac{\mathbf{L}^2}{2I} = \frac{(\mathbf{R} \times \mathbf{p})^2}{2I}$$

with $I = MR^2$.

(b) If the particle carries charge q , and the rotor is placed in a constant magnetic field $\mathbf{B}$, what is the modified Hamiltonian?

(c) What is the energy spectrum for small B ?

3. Calculate the wavelengths of the three Zeeman lines in the $3D \rightarrow 2P$ transition in hydrogen, when the latter is in a field of 10^4 gauss.

4. Consider a charged particle in a magnetic field $\mathbf{B} = (0, 0, B)$ and in a crossed electric field $\mathbf{E} = (E, 0, 0)$. Solve the eigenvalue problem. (*Hint:* The proper choice of gauge is important.)
5. Consider an electron confined to a region between two cylinders of radii a and b , ($b > a$), respectively. Inside the cylinder of radius a there is a constant magnetic field B lined up along the direction of the axis. The field vanishes outside the inner cylinder. Write down the Schrödinger equation in cylindrical coordinates. Write down the conditions that determine the energy eigenvalues. Show that the energy eigenvalues depend on the flux Φ inside the inner cylinder, even though the electron wave function is constrained to lie in the region in which B vanishes. (*Note:* For an explicit solution you need to know something about Bessel functions. For a brief summary, see Supplement 16-B) [www.wiley.com/college/gasiorowicz].

Supplement 16-A

The Aharanov–Bohm Effect

Let us return to the description of an electron of charge $-e$ and mass m_e , in a time-independent magnetic field. The system is described by the Hamiltonian

$$H = \frac{1}{2m_e} (-i\hbar \nabla + e\mathbf{A}(\mathbf{r}))^2 - e\phi(\mathbf{r}) \quad (16A-1)$$

and the time-independent Schrödinger equation reads

$$H\psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (16A-2)$$

In the absence of a vector potential the Hamiltonian reads

$$H_0 = \frac{1}{2m_e} (-i\hbar \nabla)^2 - e\phi(\mathbf{r}) \quad (16A-3)$$

and the Schrödinger equation reads

$$H_0\psi_0(\mathbf{r}) = E\psi_0(\mathbf{r}) \quad (16A-4)$$

We can show that *formally* the solution of (16A-4) and (16A-2) are related by a simple phase factor. Let us write a general expression

$$\psi(\mathbf{r}) = e^{ie\Lambda(\mathbf{r})} \psi_0(\mathbf{r}) \quad (16A-5)$$

It follows that

$$(-i\hbar \nabla + e\mathbf{A}) \psi = e\hbar \nabla \Lambda e^{ie\Lambda} \psi_0 + e^{ie\Lambda} (-i\hbar \nabla + e\mathbf{A}) \psi_0 \quad (16A-6)$$

We now observe that if we choose Λ such that

$$e\hbar \nabla \Lambda + e\mathbf{A} = 0 \quad (16A-7)$$

then

$$(-i\hbar \nabla + e\mathbf{A}) \psi = e^{ie\Lambda} (-i\hbar \nabla) \psi_0 \quad (16A-8)$$

When this is repeated, we obtain

$$H\psi = e^{ie\Lambda} H_0\psi_0 = e^{ie\Lambda} E\psi_0 = E\psi \quad (16A-9)$$

What we have just shown is that we can relate a solution of the Schrodinger equation with a vector potential $\mathbf{A}(\mathbf{r})$ to one of the Schrodinger equation *without* a vector potential by performing the *gauge transformation* shown in Eq. (16A-5). This depended, however, on our ability for find a gauge function $\Lambda(\mathbf{r})$ that satisfies Eq. (16A-7). This can only happen if

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}) = -\hbar \nabla \times \nabla \Lambda(\mathbf{r}) = 0$$

You may wonder why we bother to go through this, since if $\mathbf{B} = 0$, we did not have to write out the Schrodinger equation in the form (16A-1). The point is that in quantum mechanics it is the potentials that matter, and there are circumstances where their presence creates physical

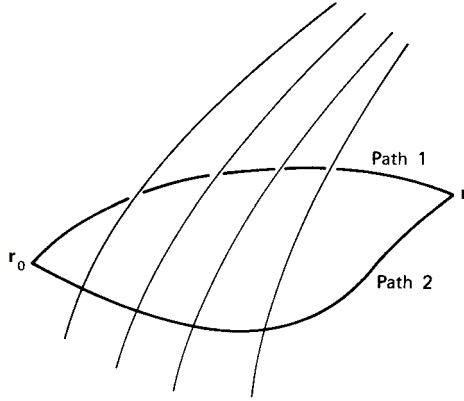


Figure 16A-1 The integrals $\int_{r_0}^r \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}'$ along path 1 and path 2 are generally not the same, since the difference is equal to the magnetic flux Φ enclosed by the closed loop.

effects even when the magnetic field vanishes. We can write out the dependence on the vector potential by noting that Eq. (16A-7) can be solved by a *line integral*,

$$\Lambda(\mathbf{r}) = -\frac{1}{\hbar} \int_P^r d\mathbf{l}' \cdot \mathbf{A}(\mathbf{r}')$$

which starts at some fixed point P and goes to $\mathbf{r}$. In terms of this (16A-5) takes the form

$$\psi(\mathbf{r}) = e^{-i(e/\hbar) \int_P^r d\mathbf{l}' \cdot \mathbf{A}(\mathbf{r}')} \psi_0(\mathbf{r}) \quad (16A-10)$$

Is the line integral in the phase independent of the path taken between P and $\mathbf{r}$? Let us consider two paths as shown in Fig. 16A-1. The difference between the line integrals is

$$\int_2 d\mathbf{l} \cdot \mathbf{A} - \int_1 d\mathbf{l} \cdot \mathbf{A} = \oint_{\text{counterclockwise}} d\mathbf{l} \cdot \mathbf{A} \quad (16A-11)$$

By Stokes' theorem we have

$$\oint d\mathbf{l} \cdot \mathbf{A} = \iint_{\text{encl. surface}} d\mathbf{S} \cdot \nabla \times \mathbf{A} = \iint_S d\mathbf{S} \cdot \mathbf{B} = \Phi \quad (16A-12)$$

This shows that for all paths that do not enclose any magnetic flux, the integral in the phase factor is the same, so that the phase factor does not depend on the path taken. The important point is that there may be situations when there is a local magnetic field, and paths that go around the flux tube are not equivalent.

In 1959 Y. Aharonov and D. Bohm pointed out a previously overlooked¹ aspect of the quantum mechanics of a charged particle in the presence of electromagnetic fields—namely, that even in a field free region, in which $\mathbf{B} = 0$, the presence in a field elsewhere can have physical consequences because $\mathbf{A} \neq 0$.

Consider, for example, a two-slit diffraction experiment carried out with electrons, as shown in Fig. 16A-2. Suppose that there is a solenoid perpendicular to the plane in which the electrons move, located as in Fig. 16A-2(a) away from the slits. The interference pattern at the screen depends on the difference in phase between the wave functions for the

¹Actually the A-B effect was discovered in an earlier paper by W. Ehrenberg and R. E. Siday (1949). Since the effects of potentials were not of central interest in the paper, their important remarks on this subject were completely overlooked.

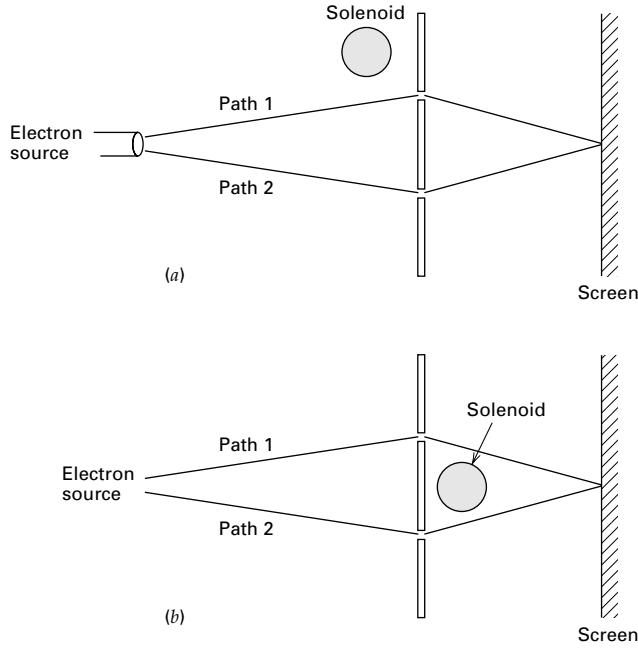


Figure 16A-2 Schematic sketch of experiment measuring shift of electron interference pattern by confined magnetic flux.

electrons passing through the two slits. In the absence of a magnetic field the interference pattern emerges from the relative phases of ψ_1 and ψ_2 in

$$\psi = \psi_1 + \psi_2 = R_1 e^{iS_1} + R_2 e^{iS_2} = e^{iS_1} (R_1 + e^{i(S_2 - S_1)} R_2) \quad (16A-13)$$

In the presence of a solenoid, (16A-13) is replaced by

$$\begin{aligned} \psi &= e^{-\frac{ie}{\hbar} \int_1 d\ell' \cdot \mathbf{A}(\mathbf{r}')} \psi_1 + e^{-\frac{ie}{\hbar} \int_2 d\ell' \cdot \mathbf{A}(\mathbf{r}')} \psi_2 \\ &= e^{-\frac{ie}{\hbar} \int_1 d\ell' \cdot \mathbf{A}(\mathbf{r}')} e^{iS_1} (R_1 + R_2 e^{i(S_2 - S_1) + ie\Phi/\hbar}) \end{aligned} \quad (16A-14)$$

If the solenoid is placed as in Fig. 16A-2(a), then no flux is enclosed by the paths of the electrons and there is no change in the interference pattern. If the flux is placed *between* the slits, as in Fig. 16A-2(b), then there is an additional contribution to the phase difference between ψ_1 and ψ_2 , so that the optical path difference is changed by a constant that depends on the enclosed flux. This has the effect of shifting the peak interference pattern from the previous center, by an amount that depends on the enclosed flux. The first experimental confirmation of the effect is due to R. G. Chambers in 1960. The definitive experiments were done by A. Tonomura and collaborators in 1980.

The A-B paper generated a certain amount of controversy, because many people believed that since only electric and magnetic fields were “physical” nothing could depend on vector potentials. As was pointed out by M. Peshkin and others, the existence of the A-B effect is intimately tied to the quantization of angular momentum, and its *absence* would raise serious questions about quantum mechanics.

Consider a particle of charge $-e$ confined to a very thin torus with radius ρ lying in the x - y plane. A solenoid with radius $a \ll \rho$ is placed along the z -axis. The magnetic field B inside the solenoid points in the positive z -direction. The vector potential in a convenient, cylindrically symmetric gauge is

$$\mathbf{A} = -\frac{1}{2} \mathbf{r} \times \mathbf{B}$$

that is,

$$\begin{aligned} A_\rho &= A_z = 0 \\ A_\phi &= \frac{1}{2} \rho B \end{aligned} \quad (16A-15)$$

The Hamiltonian operator has the form

$$\begin{aligned} & \frac{1}{2m_e} (-i\hbar \nabla + e\mathbf{A})^2 + V(\rho) \\ &= \frac{1}{2m_e} (-\hbar^2 \nabla^2 - 2ie\hbar \mathbf{A} \cdot \nabla + e^2 A^2) + V(\rho) \\ &= -\frac{\hbar^2}{2m_e} \left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \right) - \frac{ie\hbar B}{2m_e} \frac{\partial}{\partial \phi} + \frac{e^2 B^2 \rho^2}{8m_e} + V(\rho) \end{aligned} \quad (16A-16)$$

Now

$$L_z = \frac{\hbar}{i} \frac{\partial}{\partial \phi} \quad (16A-17)$$

so that the Hamiltonian operator takes the form

$$\begin{aligned} & -\frac{\hbar^2}{2m_e} \left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} \right) + \frac{L_z^2}{2m_e \rho^2} + \frac{eBL_z}{2m_e} + \frac{e^2 B^2 \rho^2}{8m_e} + V(\rho) \\ &= -\frac{\hbar^2}{2m_e} \left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} \right) + \frac{1}{2m_e \rho^2} \left(L_z + \frac{eB\rho^2}{2} \right)^2 + V(\rho) \end{aligned} \quad (16A-18)$$

Let us now take $V(\rho)$ such that the electron is confined to a torus, so narrow that we may treat ρ as *constant*. In that case the Hamiltonian becomes

$$H = \frac{1}{2m_e \rho^2} \left(L_z + \frac{e\Phi}{2\pi} \right)^2 \quad (16A-19)$$

aside from a constant. The eigenfunctions of H are eigenfunctions of L_z . With

$$L_z \psi = m\hbar \psi \quad (16A-20)$$

where $m = 0, \pm 1, \pm 2, \dots$, we find the energy eigenvalues to be

$$E = \frac{\hbar^2}{2m_e \rho^2} \left(m + \frac{e\Phi}{2\pi\hbar} \right)^2 \quad (16A-21)$$

It is clear that the energy depends on the flux, even though the electron wave function nowhere overlaps the external magnetic field. To avoid this one would have to abandon quantization of angular momentum when the x-y plane has “holes” in it, or one would have to abandon the Schrödinger equation, or the measurability of energies (or rather energy differences) in the presence of such confined fluxes. There is no question of the correctness of this effect.

There is an interesting sidelight to the A-B effect. Consider a ring made of a superconductor placed in an external magnetic field, at a temperature above the critical temperature T_c below which the material becomes superconducting. When the temperature is lowered the superconductor *expels* the magnetic field (Fig. 16A-3) except for a thin surface layer so that $\mathbf{B} = 0$ inside the material (this is the so-called *Meissner effect*),² and

²Chapter 21 in the *Feynman Lectures on Physics, Vol III* contains an excellent discussion of these macroscopic manifestations of quantum mechanics.

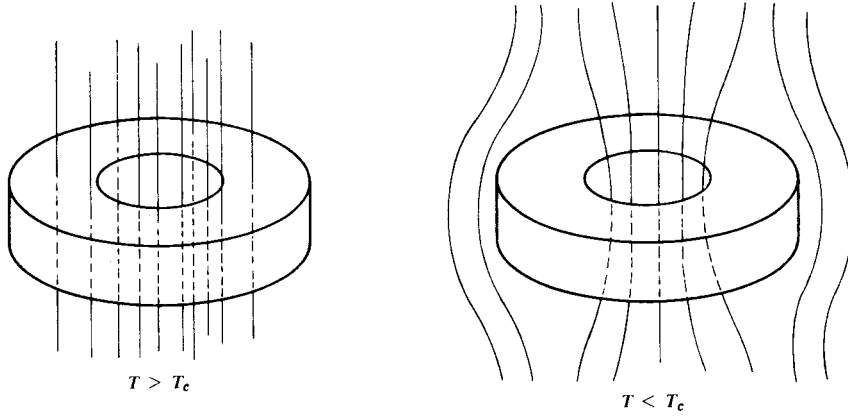


Figure 16A-3 A superconductor at temperature $T > T_c$ (the critical temperature) acts like any other metal, and magnetic flux lines can penetrate it. When the temperature is lowered until $T < T_c$, the ring becomes superconducting, and expels magnetic flux lines. Some of these become trapped inside the ring. It is the trapped flux that is found to be quantized.

magnetic flux is trapped inside the ring. The wave function of the superconductor is single-valued, and this implies that the phase factor that appears in (16A-10), when taken around a path that is inside the superconducting ring and encircles the flux region, must be unity. This implies that

$$e^{ie\Phi/\hbar} = 1 \quad (16A-22)$$

Thus the flux inside the ring is *quantized*, with

$$\frac{e\Phi}{\hbar} = 2\pi n \quad (16A-23)$$

This is almost right. The only modification that must be made is that the superconductor consists of a *condensate* of “correlated electron pairs,” so that the appropriate charge that appears in eq. (16A-23) is $2e$, and the flux quantization reads

$$\Phi = \frac{2\pi\hbar}{2e} n \quad (16A-24)$$

where n is an integer. The effect has been measured and the prediction (16A-24) was confirmed.

Supplement 16-B

A Little About Bessel Functions

The solution of the equation

$$\frac{d^2 u}{dz^2} + \frac{1}{z} \frac{du}{dz} + \left(1 - \frac{n^2}{z^2}\right) u = 0$$

with n integral, are known as Bessel functions, for the regular solutions

$$J_n(z) = \left(\frac{z}{2}\right)^n \sum_{l=0}^{\infty} \frac{(iz/2)^{2l}}{l!(n+l)!}$$

and Neumann functions for the irregular solutions

$$N_n(z) = \frac{2}{\pi} J_n(z) \log \frac{\gamma z}{2} - \frac{1}{\pi} \left(\frac{z}{2}\right)^n \sum_{l=0}^{\infty} \frac{(iz/2)^{2l}}{l!(n+l)!} a_{nl} - \frac{1}{\pi} \left(\frac{z}{2}\right)^{-n} \sum_{l=0}^{n-1} \frac{(n-l-1)!}{l!} \left(\frac{z}{2}\right)^{2l}$$
$$(\log \gamma = 0.5772 \dots) \quad a_{nl} = \left(\sum_{m=1}^l \frac{1}{m} + \sum_{m=1}^{l+n} \frac{1}{m} \right)$$

They have the asymptotic behavior

$$J_n(z) \sim \left(\frac{2}{\pi z}\right)^{1/2} \cos\left(z - \frac{n\pi}{2} - \frac{\pi}{4}\right) \left[1 + O\left(\frac{1}{z}\right)\right]$$
$$N_n(z) \sim \left(\frac{2}{\pi z}\right)^{1/2} \sin\left(z - \frac{n\pi}{2} - \frac{\pi}{4}\right) \left[1 + O\left(\frac{1}{z}\right)\right]$$

A detailed discussion of their properties may be found in any book on the special functions of mathematical physics.

Chapter 17

Radiative Decays

17-1 TRANSITION RATE

In Chapter 16 we discussed the Hamiltonian that describes the interaction of charged particles with an electromagnetic field, and studied in some detail the interaction with a constant, uniform magnetic field. In this chapter we discuss radiative transitions, with emphasis on what are usually called *electric dipole* transitions. Our starting point is eq. (16-17), with the choice of Coulomb gauge specified by

$$\nabla \cdot \mathbf{A} = 0 \quad (17-1)$$

This implies that in the operator

$$\frac{1}{2m_e} (\mathbf{p} + e\mathbf{A})^2 = \frac{1}{2m_e} (-i\hbar\nabla + e\mathbf{A})^2$$

We may write the cross term in the expansion of this as

$$\frac{1}{2m_e} (-2ie\hbar\mathbf{A} \cdot \nabla) = \frac{e}{m_e} \mathbf{A} \cdot \mathbf{p} \quad (17-2)$$

We shall be interested in the emission and absorption of *single photons*, and we will see later that this implies that we can neglect the term $e^2 A^2/2m_e$ in the Hamiltonian.

We will use time-independent perturbation theory. Writing

$$H = H_0 + \lambda V(t) \quad (17-3)$$

we will take for H_0 the Hamiltonian for the atom (e.g., the hydrogen atom), and take

$$\lambda V(t) = \frac{e}{m_e} \mathbf{A} \cdot \mathbf{p} \quad (17-4)$$

$$\mathbf{A}(\mathbf{r}, t) = \mathbf{A}_0(\mathbf{r})e^{-i\omega t} + \mathbf{A}_0^+(\mathbf{r})e^{i\omega t} \quad (17-5)$$

The perturbation is harmonic in time, and we can take over a great deal of material from Chapter 15. In particular, the first term will correspond to an increase in the energy of the system described by H_0 —that is, the *absorption* of a photon—while the second term will represent the *emission* of a photon.

Away from the source of the electromagnetic field, we find from (16-12) that

$$-\nabla^2 \mathbf{A}_0(\mathbf{r}) - \frac{\omega^2}{c^2} \mathbf{A}_0(\mathbf{r}) = 0 \quad (17-6)$$

whose solution is

$$\mathbf{A}_0(\mathbf{r}) = \mathbf{A}_0 e^{i\mathbf{k} \cdot \mathbf{r}} \quad (17-7)$$

with $\mathbf{k}^2 = \omega^2/c^2$. We chose the sign of the exponent so that the space-time dependence is $e^{\pm i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$. This may be used without loss of generality, since the direction of $\mathbf{k}$ is not specified. With this solution, the electric and magnetic fields are

$$\begin{aligned}\mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} = i\omega \mathbf{A}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} - i\omega \mathbf{A}_0^+ e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \\ \mathbf{B} &= \nabla \times \mathbf{A} = -i\mathbf{k} \times \mathbf{A}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + i\mathbf{k} \times \mathbf{A}_0^+ e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\end{aligned}\quad (17-8)$$

Note that the equation $\nabla \cdot \mathbf{E} = 0$ implies that

$$\mathbf{k} \cdot \mathbf{A}_0 = \mathbf{k} \cdot \mathbf{A}_0^+ = 0 \quad (17-9)$$

The energy density of the electromagnetic field is

$$\begin{aligned}U &= \frac{\epsilon_0}{2} \mathbf{E}^2 + \frac{1}{2\mu_0} \mathbf{B}^2 \\ &= \frac{\epsilon_0}{2} [\omega^2 (\mathbf{A}_0 \cdot \mathbf{A}_0^+ + \mathbf{A}_0^+ \cdot \mathbf{A}_0)] \\ &\quad + \frac{1}{2\mu_0} [(\mathbf{k} \times \mathbf{A}_0) \cdot (\mathbf{k} \times \mathbf{A}_0^+) + (\mathbf{k} \times \mathbf{A}_0^+) \cdot (\mathbf{k} \times \mathbf{A}_0)] \\ &\quad + \text{oscillating terms}\end{aligned}\quad (17-10)$$

If we average over time, the oscillating terms drop out. If we use (17-9), the vector identity

$$(\mathbf{k} \times \mathbf{A}_0) \cdot (\mathbf{k} \times \mathbf{A}_0^+) + (\mathbf{k} \times \mathbf{A}_0^+) \cdot (\mathbf{k} \times \mathbf{A}_0) = k^2 (\mathbf{A}_0 \cdot \mathbf{A}_0^+ + \mathbf{A}_0^+ \cdot \mathbf{A}_0)$$

and

$$\frac{1}{\epsilon_0 \mu_0} k^2 = c^2 k^2 = \omega^2$$

we end up with

$$U = \epsilon_0 \omega^2 (\mathbf{A}_0 \cdot \mathbf{A}_0^+ + \mathbf{A}_0^+ \cdot \mathbf{A}_0) \quad (17-11)$$

We may treat the electromagnetic field *semiclassically*. What this means is that we do *not* treat $\mathbf{A}$ and $\mathbf{A}^+$ as operators, but we nevertheless ascribe to the electromagnetic field energy the property that it is carried by photons of energy $\hbar\omega$. In this case, (17-11) may be rewritten as

$$U = 2\epsilon_0 \omega^2 |\mathbf{A}_0|^2 \quad (17-12)$$

If the radiation is enclosed in a volume V , then the total energy of the electromagnetic field is

$$\int_V d^3r U = 2\epsilon_0 \omega^2 |\mathbf{A}_0|^2 V \quad (17-13)$$

Let us now assume that this energy is carried by N photons, each of energy $\hbar\omega$. We may then set the right-hand side of (17-13) equal to $N\hbar\omega$. This then gives us the magnitude of $\mathbf{A}_0$.

The direction of $\mathbf{A}_0$ is specified by the polarization of the electric field. We denote it by the unit vector $\boldsymbol{\epsilon}$. It must satisfy

$$\begin{aligned}\boldsymbol{\epsilon} \cdot \boldsymbol{\epsilon} &= 1 \\ \mathbf{k} \cdot \boldsymbol{\epsilon} &= 0\end{aligned}\quad (17-14)$$

Actually there are *two* directions of polarization that are perpendicular to $\mathbf{k}$. In a semiclassical treatment, the magnitude of the vector potential comes from expressing (17-13) in

terms of photons. The net result is that for N photons of angular frequency ω the vector potential has the form

$$\mathbf{A}_0 = \epsilon \sqrt{\frac{N\hbar}{2\epsilon_0\omega V}} \quad (17-15)$$

This needs modification because there are two polarization states, which we label by λ , with $\lambda = 1, 2$. If we put in the suppressed notation, which indicates that the number of photons also depends on $\mathbf{k}$ and on λ , we write

$$\mathbf{A}(\mathbf{r}, t) = \sqrt{\frac{\hbar}{2\epsilon_0\omega V}} \sum_{\lambda=1}^2 \epsilon^{(\lambda)} [A_{\lambda}(\mathbf{k})e^{+i(\mathbf{k}\cdot\mathbf{r} - \omega t)} + A_{\lambda}^+(\mathbf{k})e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega t)}] \quad (17-16)$$

In the semiclassical approximation we have, as indicated above,

$$A_{\lambda}(\mathbf{k}) = A_{\lambda}^+(\mathbf{k}) = \sqrt{N_{\lambda}(\mathbf{k})} \quad (17-17)$$

If we treat photons quantum mechanically, as outlined in Supplement 18-A [www.wiley.com/college/gasiorowicz], the A and A^+ are operators, like the harmonic oscillator lowering and raising operators. These act on N photon states as follows:

$$\begin{aligned} A_{\lambda}(\mathbf{k}) |N_{\lambda}(\mathbf{k})\rangle &= \sqrt{N_{\lambda}(\mathbf{k})} |N_{\lambda}(\mathbf{k}) - 1\rangle \\ A_{\lambda}^+(\mathbf{k}) |N_{\lambda}(\mathbf{k})\rangle &= \sqrt{N_{\lambda}(\mathbf{k}) + 1} |N_{\lambda}(\mathbf{k}) + 1\rangle \end{aligned} \quad (17-18)$$

We see that for $N_{\lambda}(\mathbf{k}) \gg 1$ the semiclassical approximation will give the same result as the proper quantum treatment of photons.

Let us now consider a special situation. We start from an excited state of a hydrogen-like atom, $|\phi_n\rangle$, with no photons present. There is a transition that leads to a lower energy state of the atom $|\phi_m\rangle$, with the emission of a single photon. This corresponds to taking the photon states in (17-18) with $N_{\lambda} = 0$. The vector potential operative here is now

$$\mathbf{A}(\mathbf{r}, t) = \sqrt{\frac{\hbar}{2\epsilon_0\omega V}} \epsilon^{(\lambda)} e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \quad (17-19)$$

The time dependence goes into the appearance of the energy-conservation delta function [see (15-14) on p. 240]. We are therefore left with the need to evaluate the absolute square of the matrix element of the rhs in (17-4), and this must be multiplied by $2\pi/\hbar$ to get the transition rate. We therefore need to calculate

$$\frac{2\pi}{\hbar} \left| \left\langle \phi_m \left| \frac{e}{m_e} \sqrt{\frac{\hbar}{2\epsilon_0\omega V}} e^{-i\mathbf{k}\cdot\mathbf{r}} \epsilon^{(\lambda)} \cdot \mathbf{p}_{\text{op}} \right| \phi_n \right\rangle \right|^2 = \frac{\pi e^2}{m_e^2 \epsilon_0 \omega V} |\langle \phi_m | e^{-i\mathbf{k}\cdot\mathbf{r}} \epsilon^{(\lambda)} \cdot \mathbf{p}_{\text{op}} | \phi_n \rangle|^2 \quad (17-20)$$

17-2 CALCULATION OF THE MATRIX ELEMENT

Our task is to calculate $\langle \phi_m | e^{-i\mathbf{k}\cdot\mathbf{r}} \epsilon^{(\lambda)} \cdot \mathbf{p}_{\text{op}} | \phi_n \rangle$. We begin by estimating the magnitude of the matrix element. For a typical atomic state,

$$\epsilon \cdot \mathbf{p} \approx |p| \approx Z m_e c \alpha \quad (17-21)$$

This is not enough, since there is an oscillating factor $e^{-i\mathbf{k}\cdot\mathbf{r}}$ appears, and that can change the estimate considerably. With

$$r \approx \frac{\hbar}{m_e Z \alpha c}$$

and

$$|k| = \frac{\hbar\omega}{\hbar c} \approx \frac{m_e(Z\alpha c)^2}{2\hbar c} \approx \frac{m_e c(Z\alpha)^2}{2\hbar}$$

we get $kr = Z\alpha/2$. This means that for $Z\alpha \ll 1$ we can set the exponential factor equal to unity. It simplifies matters that in the expansion

$$e^{-i\mathbf{k}\cdot\mathbf{r}} = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} (\mathbf{k} \cdot \mathbf{r})^n$$

the successive terms are estimated to decrease as $Z\alpha$.

The leading term may be rewritten as follows:

$$\begin{aligned} \langle \phi_m | \boldsymbol{\epsilon}^{(\lambda)} \cdot \mathbf{p}_{\text{op}} | \phi_n \rangle &= \boldsymbol{\epsilon}^{(\lambda)} \cdot \left\langle \phi_m \left| m_e \frac{d\mathbf{r}}{dt} \right| \phi_n \right\rangle \\ &= m_e \boldsymbol{\epsilon}^{(\lambda)} \cdot \left\langle \phi_m \left| \frac{i}{\hbar} [H_0, \mathbf{r}] \right| \phi_n \right\rangle \\ &= im_e \boldsymbol{\epsilon}^{(\lambda)} \cdot \left\{ \frac{E_m^{(0)} - E_n^{(0)}}{\hbar} \langle \phi_m | \mathbf{r} | \phi_n \rangle \right\} \\ &= -im_e \omega \langle \phi_m | \boldsymbol{\epsilon}^{(\lambda)} \cdot \mathbf{r} | \phi_n \rangle \end{aligned} \quad (17-22)$$

In the final analysis, we need to calculate the matrix elements of the operator $\mathbf{r}$. This arises from the approximation in which we replaced the oscillating factor by unity. The approximation is called the *electric dipole approximation*. A justification for this name comes from the fact that in this approximation, with (17-8),

$$\frac{e}{m_e} \mathbf{A}(\mathbf{r}, t) \cdot \mathbf{p} = e\mathbf{E}(\mathbf{r}, t) \cdot \mathbf{r}$$

and this, in fact, is the potential energy of a dipole, with moment $\mathbf{d} = -e\mathbf{r}$, in an electric field $\mathbf{E}$.

For hydrogenlike atoms, the matrix element in (17-22) involves an initial state $|\phi_n\rangle$, which we shall characterize by the quantum numbers (n_i, l_i, m_i) and a final state $|\phi_m\rangle$ with quantum numbers (n_f, l_f, m_f) . We will therefore want to calculate

$$\begin{aligned} \langle \phi_m | \boldsymbol{\epsilon}^{(\lambda)} \cdot \mathbf{r} | \phi_n \rangle &= \int_0^\infty r^2 dr \int d\Omega R_{n_f l_f}^*(r) Y_{l_f m_f}^*(\theta, \varphi) \boldsymbol{\epsilon}^{(\lambda)} \cdot \mathbf{r} R_{n_i l_i}(r) Y_{l_i m_i}(\theta, \varphi) \\ &= \int_0^\infty r^3 dr R_{n_f l_f}^*(r) R_{n_i l_i}(r) \times \int d\Omega Y_{l_f m_f}^*(\theta, \varphi) \boldsymbol{\epsilon}^{(\lambda)} \cdot \hat{\mathbf{r}} Y_{l_i m_i}(\theta, \varphi) \end{aligned} \quad (17-23)$$

The radial integral will be discussed for a special case in a later section. At this point we concentrate on the angular integral.

17-3 ANGULAR INTEGRATION AND SELECTION RULES

The angular integral involves the quantity

$$\boldsymbol{\epsilon}^{(\lambda)} \cdot \hat{\mathbf{r}} = \epsilon_x^{(\lambda)} \sin \theta \cos \varphi + \epsilon_y^{(\lambda)} \sin \theta \sin \varphi + \epsilon_z^{(\lambda)} \cos \theta \quad (17-24)$$

In what follows we leave off the λ superscript for simplicity. It will reappear at the end. A look at the table of spherical harmonics shows that

$$\begin{aligned}\cos \theta &= \sqrt{\frac{4\pi}{3}} Y_{10}(\theta, \varphi) \\ \sin \theta e^{\pm i\varphi} &= \mp \sqrt{\frac{8\pi}{3}} Y_{1,\pm 1}(\theta, \varphi)\end{aligned}\quad (17-25)$$

so that

$$\boldsymbol{\varepsilon} \cdot \hat{\mathbf{r}} = \sqrt{\frac{4\pi}{3}} \left(\varepsilon_z Y_{10} + \frac{-\varepsilon_x + i\varepsilon_y}{\sqrt{2}} Y_{1,1} + \frac{\varepsilon_x + i\varepsilon_y}{\sqrt{2}} Y_{1,-1} \right) \quad (17-26)$$

The angular integration in (17-23) therefore involves

$$\int d\Omega Y_{l_f m_f}^*(\theta, \varphi) Y_{1,m}(\theta, \varphi) Y_{l_i m_i}(\theta, \varphi) \quad (17-27)$$

Let us first consider the azimuthal integration. It involves

$$\int_0^{2\pi} d\varphi e^{-im_f \varphi} e^{im \varphi} e^{im_i \varphi} = 2\pi \delta_{m=m_f-m_i} \quad (17-28)$$

The Kronecker delta gives us the first *selection rule*. It requires that in dipole transitions we must have

$$m_f - m_i = m = 1, 0, -1 \quad (17-29)$$

We saw this selection rule earlier when we discussed the transitions in the Zeeman effect. Actually we can say more: If the photon momentum direction $\mathbf{k}$ is chosen to define the z -axis, then the *transversality condition* in (17-14) implies that the $\boldsymbol{\varepsilon}$ must lie in the x - y plane. This means that m above can only be equal to ± 1 , and our selection rule states that

$$m_f - m_i = \pm 1 \quad (17-30)$$

As a special case, consider the final state to be the ground state, with $l_f = m_f = 0$. In that case, $m = m_i$. For example if $m_i = 1$, then $m = -1$, and hence the polarization for the radiation is $(\varepsilon_x + i\varepsilon_y)/\sqrt{2}$. The implication is that if the atom in the initial state is polarized along the z -axis with $m_i = 1$, then in a decay to a state with zero angular momentum, *angular momentum conservation*, specifically that of the z -component, requires that the outgoing photon carry off the initial angular momentum. Thus the photon must have its spin aligned along the positive z -axis, its direction of motion. The component of spin of a particle along its direction of motion is called the *helicity*. In the special case under discussion, the helicity of the photon must be positive. This is equivalent to the statement that the photon has a left-circular polarization, which is what the term $(\varepsilon_x + i\varepsilon_y)/\sqrt{2}$ implies.

The θ integration gives rise to another selection rule. Again, in the special case that $l_f = 0$, the final wave (angular) function $Y_{00} = 1/\sqrt{4\pi}$ leads to the integral

$$\int d\Omega \frac{1}{\sqrt{4\pi}} Y_{1,m}(\theta, \varphi) Y_{l_i m_i}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}} \delta_{l_i,1} \delta_{m_i,-m} \quad (17-31)$$

which implies that the initial state *must have* $l_i = 1$. In hydrogen, the dominant transition to the ground state will be $np \rightarrow 1s$.

More generally, when neither the initial or final l vanish, we still get a selection rule. We do not have the machinery at our disposal to work this out, but the result is simply a

consequence of the rules of addition of angular momentum and of angular momentum conservation. The initial angular momentum is l_i . In the final state the angular momentum is the sum of that of the atom l_f and, *in the dipole approximation*, the angular momentum of the photon reflected in the wave function $Y_{1,m}$. Our rules of addition of angular momentum, stated in Chapter 10, tell us that the total angular momentum in the final state is one of $l_f + 1, l_f, |l_f - 1|$. This must be equal to l_i . We can thus rewrite the *selection rule for the orbital angular momentum* as

$$l_f = l_i + 1, l_i, |l_i - 1|$$

This is the general form of the *electric dipole selection rule*

$$\Delta l = \pm 1, 0 \quad (17-32)$$

Note that if the initial and final states both have $l = 0$, then the integral has the form $\int d\Omega Y_{1m}(\theta, \varphi) = 0$. Thus, *there are no zero-zero transitions*. There is a further constraint that comes from parity conservation. Since $\mathbf{r}$ is *odd* under reflections ($x \rightarrow -x$ and so on), there is an additional selection rule that states

$$\text{The atomic state must change parity} \quad (17-33)$$

Since the parity of a spherical harmonic Y_{lm} is $(-1)^l$, this implies that the l -value of the state must actually change. For example, $3p \rightarrow 2p$ transitions, allowed by the rule (17-32), as stated, cannot actually take place. Thus (17-32) should really be replaced by

$$\Delta l = \pm 1 \quad (17-32')$$

To the extent that the only coupling of the electromagnetic field to the atoms is through the perturbation

$$\lambda V(t) = \frac{e}{m_e} \mathbf{A}(\mathbf{r}, t) \cdot \mathbf{p}$$

there is no spin dependence in it, and thus if the initial state is an “up” state, then the final state will also be an “up” state. In other words, the spin cannot flip over in a transition. This is equivalent to a selection rule

$$\Delta S = 0 \quad (17-34)$$

as mentioned earlier in connection with the spectrum of helium.

The selection rules stated above are not absolute. The conservation laws of angular momentum and parity (for electromagnetic processes) are absolute, but (17-32') is only approximately true. Transitions between states that involve changes of l larger than one cannot take place through the electric dipole mechanism, but they can occur as long as there is a nonvanishing matrix element.

$$\langle \phi_f | e^{-i\mathbf{k} \cdot \mathbf{r}} \boldsymbol{\varepsilon} \cdot \mathbf{p}_{\text{op}} | \phi_i \rangle$$

For $\Delta l = 2$, the first power of $\mathbf{k} \cdot \mathbf{r}$ in the expansion of the exponential will give a nonvanishing contribution. We may write

$$\begin{aligned} \mathbf{k} \cdot \mathbf{r} \boldsymbol{\varepsilon} \cdot \mathbf{p} &= \frac{1}{2} (\boldsymbol{\varepsilon} \cdot \mathbf{p} \mathbf{k} \cdot \mathbf{r} + \boldsymbol{\varepsilon} \cdot \mathbf{r} \mathbf{p} \cdot \mathbf{k}) + \frac{1}{2} (\boldsymbol{\varepsilon} \cdot \mathbf{p} \mathbf{k} \cdot \mathbf{r} - \boldsymbol{\varepsilon} \cdot \mathbf{r} \mathbf{p} \cdot \mathbf{k}) \\ &= \frac{1}{2} (\boldsymbol{\varepsilon} \cdot \mathbf{p} \mathbf{k} \cdot \mathbf{r} + \boldsymbol{\varepsilon} \cdot \mathbf{r} \mathbf{p} \cdot \mathbf{k}) + \frac{1}{2} (\mathbf{k} \times \boldsymbol{\varepsilon}) \cdot (\mathbf{r} \times \mathbf{p}) \end{aligned} \quad (17-35)$$

The first of these terms is called an *electric quadrupole* term. The second is clearly related to an $\mathbf{L} \cdot \mathbf{B}$ term, and it is called a *magnetic dipole* term. For these transitions, whose matrix elements we estimated in the material following (17-21) to be $Z\alpha$ times smaller than

the electric dipole terms, we will have $\Delta l = 2$ and $\Delta l = 0$, respectively. Since the operators in (17-35) are even, there will be *no parity change* in the atomic states taking part in these transitions. These matrix elements are smaller than the leading ones, but when electric dipole transitions are forbidden, then these are the ones that matter. For example, transitions $3d \rightarrow 1s$, forbidden for dipole transitions, can go via the electric quadrupole mechanism. Actually, it turns out that a two-step dipole transition $3d \rightarrow 2p$, $2p \rightarrow 1s$ is much more probable than a direct quadrupole transition, even though in the two-step process two photons are emitted in succession.

The spin selection rule $\Delta S = 0$ is also not sacred. In addition to the coupling in (17-4) there is the coupling discussed in connection with the anomalous Zeeman effect,

$$\lambda V(t) = \frac{eg}{2m_e} \mathbf{S} \cdot \mathbf{B}(\mathbf{r}, t) \quad (17-36)$$

The matrix elements for the $\Delta S \neq 0$ transition-inducing process can be estimated. We compare it with the electric dipole matrix element

$$\frac{\frac{eg}{2m_e} \hbar |\mathbf{k} \times \boldsymbol{\varepsilon}|}{\frac{2e}{m_e} |\mathbf{p} \cdot \boldsymbol{\varepsilon}|} \approx \frac{g}{4} \frac{|\mathbf{k}|}{|\mathbf{p}|} \approx \frac{\hbar \omega}{|\mathbf{p}|c} \approx \frac{m_e c^2 (Z\alpha)^2}{m_e c^2 (Z\alpha)} = Z\alpha \quad (17-37)$$

We see that it is suppressed, just like the magnetic dipole matrix element, which it strongly resembles.

As an example where the coupling (17-36) plays an important role, we consider a process important in low-energy nuclear physics, the *photodisintegration of the deuteron*

$$\gamma + d \rightarrow p + n$$

The deuteron is a weakly bound state of a proton and a neutron. The lowest bound state in any sensible potential will be a state of orbital angular momentum zero. Since the neutron and the proton both have spin $1/2$, the deuteron can be a singlet state or a triplet state. Experiments show that the total angular momentum is $J = 1$, so that deuteron is a ${}^3S_1$ state. An electric dipole transition must lead to a final ${}^3P$ state, since $\Delta l = 1$ and the spin does not change. In general, particles will be likely to be in a relative angular momentum L state only if

$$|\mathbf{p}| a \geq \hbar L \quad (17-38)$$

where $|\mathbf{p}|$ is the magnitude of the relative momentum and a represents the range of the forces that involve the two particles. For the energies of the magnitude of the deuteron binding energy this strongly supports the $L = 0$ assignment. Let us consider the interaction analogous to (17-36):

$$-\frac{e}{2M} (g_p \mathbf{S}_p + g_n \mathbf{S}_n) \cdot \mathbf{B} \quad (17-39)$$

where M is the nucleon mass ($M_p \cong M_n$) and g_p and g_n are the gyromagnetic ratios for the proton and neutron, respectively. This interaction can lead to a transition between the ${}^3S_1$ bound state and the unbound ${}^1S_0$ n - p state. The interaction (17-39) may be rewritten in the form

$$-\frac{e}{2M} \left[\frac{g_p + g_n}{2} (\mathbf{S}_p + \mathbf{S}_n) + \frac{g_p - g_n}{2} (\mathbf{S}_p - \mathbf{S}_n) \right] \cdot \mathbf{B} \quad (17-40)$$

The initial spin state is a total $S = 1$ state, and we know from our discussion of the addition of two spins in Chapter 10 that this state is symmetric in the n and p spinors. The final state is an $S = 0$ state, antisymmetric in the n and p spinors. Thus the first term, consisting

of the sum of the spin operators, is symmetric, and it cannot therefore act on a symmetric state and lead to an antisymmetric one. Thus only the second term contributes. The coefficient in front is actually quite large since $g_p = 5.56$ and $g_n = -3.81$.

There is one selection rule that is sacred, and that is the one forbidding zero-zero transitions (referring to *total* angular momentum $j = 0$) in one-photon processes. A general way of arguing the absoluteness of this selection rule is the following: The matrix element, a scalar quantity, must involve the photon polarization linearly, and must therefore be of the form $\boldsymbol{\varepsilon} \cdot \mathbf{V}$, where $\mathbf{V}$ is some vector that enters into the problem. If the initial and final states are $j = 0$ states—that is, have no directionality associated with them—then the only vector is $\mathbf{k}$, the photon momentum. However, $\boldsymbol{\varepsilon} \cdot \mathbf{k} = 0$, so that there is no way of constructing a matrix element. It must therefore not exist.¹

17-4 THE $2p \rightarrow 1s$ TRANSITION

Let us now specialize to the transition $2p \rightarrow 1s$ in (17-23). We need to evaluate the radial integral

$$\begin{aligned} \int_0^\infty dr r^3 R_{10}^*(r) R_{21}(r) &= \int_0^\infty dr r^3 \left[2 \left(\frac{Z}{a_0} \right)^{3/2} e^{-Zr/a_0} \right] \left[\frac{1}{\sqrt{24}} \left(\frac{Z}{a_0} \right)^{5/2} r e^{-Zr/2a_0} \right] \\ &= \frac{1}{\sqrt{6}} \left(\frac{Z}{a_0} \right)^4 \int_0^\infty dr r^4 e^{-3Zr/2a_0} = \frac{1}{\sqrt{6}} \left(\frac{Z}{a_0} \right)^4 \left(\frac{2a_0}{3Z} \right)^5 \int_0^\infty dx x^4 e^{-x} \quad (17-41) \\ &= \frac{24}{\sqrt{6}} \left(\frac{2}{3} \right)^5 Z^{-1} a_0 \end{aligned}$$

The angular integral is

$$\begin{aligned} d\Omega Y_{0,0}^* \boldsymbol{\varepsilon} \cdot \hat{\mathbf{r}} Y_{1,m} &= \frac{1}{\sqrt{4\pi}} \int d\Omega \sqrt{\frac{4\pi}{3}} \left(\varepsilon_z Y_{1,0} + \frac{-\varepsilon_x + i\varepsilon_y}{\sqrt{2}} Y_{1,1} + \frac{\varepsilon_x + i\varepsilon_y}{\sqrt{2}} Y_{1,-1} \right) Y_{1,m} \\ &= \frac{1}{\sqrt{3}} \left(\varepsilon_z \delta_{m,0} + \frac{-\varepsilon_x + i\varepsilon_y}{\sqrt{2}} \delta_{m,-1} + \frac{\varepsilon_x + i\varepsilon_y}{\sqrt{2}} \delta_{m,1} \right) \quad (17-42) \end{aligned}$$

The absolute square of the product of (17-41) and (17-42) is

$$96 \left(\frac{2}{3} \right)^{10} \left(\frac{a_0}{Z} \right)^2 \times \frac{1}{3} \left[\delta_{m,0} \varepsilon_z^2 + \frac{1}{2} (\delta_{m,1} + \delta_{m,-1}) (\varepsilon_x^2 + \varepsilon_y^2) \right] \quad (17-43)$$

To get the rate, we must multiply by $2\pi/\hbar$ and the phase space factor. Combining equation (17-20), (17-43), and (15-29) we get

$$d\Gamma = \frac{\pi e^2}{m_e^2 \varepsilon_0 \omega V} m_e^2 \omega^2 \left[96 \left(\frac{2}{3} \right)^{10} \left(\frac{a_0}{Z} \right)^2 \right] \frac{1}{3} \left[\delta_{m,0} \varepsilon_z^2 + \frac{1}{2} (\delta_{m,1} + \delta_{m,-1}) (\varepsilon_x^2 + \varepsilon_y^2) \right] \frac{V}{(2\pi)^3} \frac{\omega^2}{\hbar c^3} d\Omega_{\mathbf{k}}$$

The fact that the factors V cancel is always a good check on the calculation. The result is

$$d\Gamma = \frac{\alpha}{2\pi} \frac{\omega^3}{c^2} \left(\frac{a_0}{Z} \right)^2 \frac{2^{15}}{3^{10}} \left[\delta_{m,0} \varepsilon_z^2 + \frac{1}{2} (\delta_{m,1} + \delta_{m,-1}) (\varepsilon_x^2 + \varepsilon_y^2) \right] d\Omega_{\mathbf{k}} \quad (17-44)$$

¹The relation $\boldsymbol{\varepsilon} \cdot \mathbf{k} = 0$ is independent of the choice of gauge, and is a statement about the transversality of the electromagnetic field.

For the $2p \rightarrow 1s$ transition, the angular frequency of the photon emitted is

$$\omega = \frac{1}{\hbar} \left[\frac{1}{2} m_e c^2 (Z\alpha)^2 \left(1 - \frac{1}{4} \right) \right] = \frac{3}{8} \frac{m_e c^2}{\hbar} (Z\alpha)^2 \quad (17-45)$$

The differential rate needs to be integrated over the photon directions. This is not trivial, since the polarization vector $\boldsymbol{\epsilon}$ is constrained to be perpendicular to the photon direction $\hat{\mathbf{k}}$. The result is simplified if the initial $l = 1$ state of the excited atom is *unaligned*—that is, if all three m states ($m = 1, 0, -1$) are equally populated. In that case the rate, *averaged* over the three initial m states, is

$$d\Gamma = \frac{1}{3} \sum_{m=-1}^1 d\Gamma_m \quad (17-46)$$

Now

$$\sum_{m=-1}^1 \left[\delta_{m0} \epsilon_z^2 + \frac{1}{2} (\delta_{m1} + \delta_{m,-1}) (\epsilon_x^2 + \epsilon_y^2) \right] = \epsilon_z^2 + \epsilon_x^2 + \epsilon_y^2 = 1 \quad (17-47)$$

so that for this case, the integrand becomes independent of the photon direction. The result should also be multiplied by 2. As noted before, there are two polarization states for the photon, denoted by the (suppressed) label λ . Since we are detecting both polarization states, we sum over λ . In the case under consideration nothing depends on λ , so that we merely double our result. More generally we work with

$$\sum_{\lambda=1}^2 |\langle \phi_m | e^{i\mathbf{k}\cdot\mathbf{r}} \boldsymbol{\epsilon}^{(\lambda)} \cdot \mathbf{r} | \phi_n \rangle|^2 \quad (17-48)$$

If we insert the factor of 2, integrate over $d\Omega$ (which here gives 4π), and use (17-45), we end up with

$$\Gamma(2p \rightarrow 1s) = \left(\frac{2}{3} \right)^8 \alpha^5 \frac{m_e c^2}{\hbar} Z^4 = 0.62 \times 10^9 Z^4 \text{s}^{-1} \quad (17-49)$$

This result differs by a factor of about 25 from our initial rough estimate, which shows that detailed calculations are necessary for obtaining the numerical coefficients. Nevertheless, dimensional considerations and a proper counting of powers of α do give us an order of magnitude guidance to the size of physical quantities in atomic physics.

It is worthwhile to notice that we can obtain the intensity of the radiation by multiplying the rate by the photon energy. We obtain

$$\begin{aligned} \frac{dI(\omega)}{d\Omega_{\mathbf{k}}} &= \hbar\omega \frac{\pi e^2}{m_e^2 \epsilon_0 \omega V} m_e^2 \omega^2 \sum_{\lambda=1}^2 |\langle \boldsymbol{\epsilon}^{(\lambda)} \cdot \mathbf{r} \rangle|^2 \frac{V}{(2\pi)^3} \frac{\omega^2}{c^3} \\ &= \frac{1}{24} \frac{e^2}{4\pi\epsilon_0} \frac{\omega^4}{c^3} \sum_{\lambda=1}^2 |\langle \boldsymbol{\epsilon}^{(\lambda)} \cdot \mathbf{r} \rangle|^2 \end{aligned} \quad (17-50)$$

This, however, is just the *classical* formula for the intensity of light emitted by an oscillating dipole, of dipole moment

$$\mathbf{d} = e \langle \phi_m | \mathbf{r} | \phi_n \rangle e^{-i\omega t} \quad (17-51)$$

This is yet another illustration of the correspondence principle. In this calculation we paid no attention to spin; this is relegated to Supplement 17-A [www.wiley.com/college/gasiorowicz].

PROBLEMS

1. Work out the part of the matrix element in deuteron photodisintegration that only involves the spin; that is, calculate

$$X_{\text{singlet}}^* \left[\frac{g_p - g_n}{2} (\mathbf{S}_p - \mathbf{S}_n) \cdot \mathbf{B} \right] X_{\text{triplet}}$$

where

$$X_{\text{singlet}} = \frac{1}{\sqrt{2}} (\chi_+^{(p)} \chi_-^{(n)} - \chi_-^{(p)} \chi_+^{(n)})$$

and

$$X_{\text{triplet}} = (\chi_+^{(p)} \chi_+^{(n)}, \frac{1}{\sqrt{2}} (\chi_+^{(p)} \chi_-^{(n)} + \chi_-^{(p)} \chi_+^{(n)}), \chi_-^{(p)} \chi_-^{(n)})$$

For $\mathbf{B}$ use $\nabla \times \mathbf{A}$ and (17-19).

2. Show that the first term in (17-35) leads to $\Delta l = 2$ transitions, by evaluating the term between a state $Y_{lm}(\theta, \varphi)$ and the ground state ($l = 0$).
3. Consider a transition from the first excited state in hydrogen to the ground state. Assume that the excited state is *polarized*; that is, it is in a state $m = 1$. Work out $d\Gamma$ for this case.
4. Calculate the $2p \rightarrow 1s$ transition rate for a harmonic oscillator. In this case we are given that the energy eigenvalues are $E = \hbar\omega(n + \frac{3}{2})$ and the energy quantum number $n = 2n_r + l$; $n_r = 0, 1, 2, \dots$; $\ell = 0, 1, 2, \dots$.

Supplement 17-A

Spin and Intensity Rules

The inclusion of spin does not change things very much. It is true that the initial states and the final states can each be in an “up” or “down” spin state, but since the interaction in atomic transitions is spin independent, only “up” $\rightarrow$ “up” and “down” $\rightarrow$ “down” transitions are allowed. Hence the transition rates will not only be independent of m_l (as we saw in Chapter 17) but also of m_s , and hence, m_j . With the inclusion of spin-orbit coupling, there will be small (on the scale of the $2p - 1s$ energy difference) level splittings. For example, the $n = 1$ and $n = 2$ level structure is changed, as shown in Fig. 17-A1. The spectral line corresponding to the transition $2p \rightarrow 1s$ is split into two lines, $2^2P_{3/2} \rightarrow 1^2S_{1/2}$ and $2^2P_{1/2} \rightarrow 1^2S_{1/2}$. For the split states, the radial integral and the phase space are almost unchanged, and hence *the ratio of the intensity of the two lines can be determined from the angular parts of the integral alone—that is, purely from angular momentum considerations.*

The following table lists the wave functions for the states in question.

J	m_j	Odd Parity $l = 1$	Even Parity $l = 0$
3/2	3/2	$Y_{11}\chi_+$	—
3/2	1/2	$\sqrt{2/3} Y_{10}\chi_+ + \sqrt{1/3} Y_{11}\chi_-$	—
3/2	-1/2	$\sqrt{1/3} Y_{1,-1}\chi_+ + \sqrt{2/3} Y_{10}\chi_-$	—
3/2	-3/2	$Y_{1,-1}\chi_-$	—
1/2	1/2	$\sqrt{1/3} Y_{10}\chi_+ - \sqrt{2/3} Y_{11}\chi_-$	$Y_{00}\chi_+$
1/2	-1/2	$\sqrt{2/3} Y_{1,-1}\chi_+ - \sqrt{1/3} Y_{10}\chi_-$	$Y_{00}\chi_-$

In the squares of the matrix elements, the radial parts are common to all of them. Thus, in considering the rates for $P_{3/2} \rightarrow S_{1/2}$ we must add the squares of the transition matrix elements for $m_j = 3/2 \rightarrow m_j = 1/2$, $m_j = 3/2 \rightarrow m_j = -1/2$, $\dots$, $m_j = -3/2 \rightarrow m_j = -1/2$, while the rate for the $P_{1/2} \rightarrow S_{1/2}$ involves the sum of the squares of the matrix elements for $m_j = 1/2 \rightarrow m_j = 1/2$, $\dots$, $m_j = -1/2 \rightarrow m_j = -1/2$. This can be done directly by techniques that are quite sophisticated and beyond the scope of this book. One can, however, work out these quantities in detail, using the fact that the spin wave functions are orthonormal.

$$\begin{array}{ll}
 \frac{P_{3/2} \rightarrow S_{1/2}}{m_j = 3/2 \rightarrow m_j = 1/2} & |\langle Y_{11} | \mathbf{r} \cdot \boldsymbol{\epsilon} | Y_{00} \rangle|^2 = C \\
 3/2 \rightarrow -1/2 & 0 \quad \text{since } \chi_+^* \chi_- = 0 \\
 1/2 \rightarrow 1/2 & |\langle \sqrt{2/3} Y_{10} | \mathbf{r} \cdot \boldsymbol{\epsilon} | Y_{00} \rangle|^2 = 0 \quad (\Delta m = 0) \\
 1/2 \rightarrow -1/2 & |\langle \sqrt{1/3} Y_{11} | \mathbf{r} \cdot \boldsymbol{\epsilon} | Y_{00} \rangle|^2 = C/3 \\
 -1/2 \rightarrow 1/2 & |\langle \sqrt{1/3} Y_{1,-1} | \mathbf{r} \cdot \boldsymbol{\epsilon} | Y_{00} \rangle|^2 = C/3 \\
 -1/2 \rightarrow -1/2 & |\langle \sqrt{2/3} Y_{10} | \mathbf{r} \cdot \boldsymbol{\epsilon} | Y_{00} \rangle|^2 = 0 \quad (\Delta m = 0) \\
 -3/2 \rightarrow 1/2 & 0 \\
 -3/2 \rightarrow -1/2 & |\langle Y_{1,-1} | \mathbf{r} \cdot \boldsymbol{\epsilon} | Y_{00} \rangle|^2 = C
 \end{array} \tag{17A-1}$$

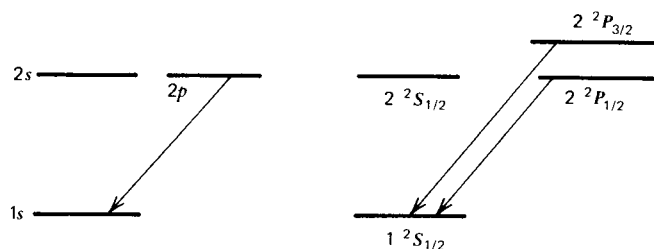


Figure 17A-1 The splitting of the $2p - 1s$ spectral line by spin-orbit coupling.

If we sum the terms we get

$$\sum R = \frac{8C}{3} \quad (17A-2)$$

Similarly,

$$\begin{array}{ll} \frac{P_{1/2} \rightarrow S_{1/2}}{m_j = 1/2 \rightarrow m_j = 1/2} & |\langle \sqrt{1/3} Y_{10} | \mathbf{e} \cdot \mathbf{r} | Y_{00} \rangle|^2 = 0 \\ 1/2 \rightarrow -1/2 & |\langle -\sqrt{2/3} Y_{11} | \mathbf{e} \cdot \mathbf{r} | Y_{00} \rangle|^2 = 2C/3 \\ -1/2 \rightarrow 1/2 & |\langle \sqrt{2/3} Y_{1,-1} | \mathbf{e} \cdot \mathbf{r} | Y_{00} \rangle|^2 = 2C/3 \\ -1/2 \rightarrow -1/2 & |\langle -\sqrt{1/3} Y_{10} | \mathbf{e} \cdot \mathbf{r} | Y_{00} \rangle|^2 = 0 \end{array} \quad (17A-3)$$

Again,

$$\sum R = \frac{4C}{3} \quad (17A-4)$$

Thus the ratio of the intensities is

$$\frac{R(P_{3/2} \rightarrow S_{1/2})}{R(P_{1/2} \rightarrow S_{1/2})} = \frac{8C/3}{4C/3} = 2 \quad (17A-5)$$

The reason for *summing* over all the initial states is that when the atom is excited, all the p -levels are equally occupied, since their energy difference is so tiny compared to the $2p - 1s$ energy difference. We also sum over all the final states if we perform an experiment that does not discriminate between them, as is the case for a spectroscopic measurement. In our calculation of the $2p \rightarrow 1s$ transition rate, we *averaged* over the initial m -states. There we were concerned with the problem of asking: "If we have N atoms in the $2p$ states, how many will decay per second?" The averaging came about because of the fact that under most circumstances, when N atoms are excited, about $N/3$ go into each one of the $m = 1, 0, -1$ states. Here, the fact that there are more levels in the $P_{3/2}$ state than there are in the $P_{1/2}$ state is relevant. There will be altogether six levels, (four with $j = 3/2$ and two with $j = 1/2$), and there will be on the average $N/6$ atoms in each of the states. The fact that there are more atoms in the $j = 3/2$ subset of levels just means that more atoms decay, and that therefore the intensity will be larger.

Chapter 18

Selected Topics on Radiation

In our discussion of radiative transitions, which involved the emission or absorption of a single photon, it was possible to treat the electromagnetic field semiclassically. We were able to treat the vector potential $\mathbf{A}(\mathbf{r}, t)$ as an external field, though its magnitude was adjusted so that the total energy associated with the electromagnetic field corresponded to that carried by a single photon, as seen in Eq. (17-16). In this chapter we will be dealing with *lasers*, and this means that we will be dealing with multiple-photon states. To do this properly we must treat the electromagnetic field quantum mechanically. A simplified version of this is outlined in Supplement 18-A. The result is that we treat the electric field as an *operator* in the form

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}, \lambda} \sqrt{\frac{\hbar \omega}{2\epsilon_0}} \epsilon_{\lambda}(\mathbf{k}) [-iA_{\lambda}(\mathbf{k})e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + iA_{\lambda}^{\dagger}(\mathbf{k})e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}] \quad (18-1)$$

Here $A_{\lambda}(\mathbf{k})$ and $A_{\lambda}^{\dagger}(\mathbf{k})$ are annihilation and creation operators for a photon with wave number $\mathbf{k}$ and polarization λ . They act exactly as the energy lowering and raising operators for the simple harmonic oscillator, as discussed in Chapter 6. In particular, the electric field operator acting on a state containing N identical photons (say, all with the attributes $(\mathbf{k}, \lambda)$) will be enhanced compared with the one-photon case by a factor $\sqrt{N(\mathbf{k}, \lambda) + 1}$, and the square of the matrix element will contain a factor $N(\mathbf{k}, \lambda) + 1$. This means that when a system emits a photon of a certain kind, and N photons of that kind are present, then the emission rate is proportional to $N + 1$. The part involving the factor 1 in this is described as *spontaneous emission*, while the factor proportional to N is described as *induced emission*. We saw this explicitly in Supplement 1-A [www.wiley.com/college/gasiorowicz]. Induced emission plays the central role in the operation of the *laser*.

18-1 LASERS

Induced emission finds its most dramatic technological application in the production of coherent, monochromatic, highly directional electromagnetic radiation by means of *light amplification through stimulated emission* in a device called the *laser*.¹ The basic components of a laser are the following:

1. A lasing medium in which there are at least two energy levels, separated by an energy gap, such that the atoms in the upper level can make a transition that is stimulated by the presence of photons of the right frequency.
2. Some mechanism for repopulating the upper level for repeated operation.
3. A suitable cavity in which the stimulating photons can be contained and that also contains the lasing medium.

¹The phenomenon holds for radiation of all frequencies and was indeed first studied in the microwave region. The device was called the Maser.

Conditions for the Operation of a Laser

Consider a material in which we focus on two energy levels with energies E_1 and E_0 and $E_1 > E_0$. We will label the number of photons purely by their frequency, suppressing the direction of $\mathbf{k}$ and the polarization λ , since we are talking about identical photons. The number of photons $N(\nu)$ will change with time. The number will increase, because of stimulated as well as spontaneous emission. If the number of emitting atoms in the upper level is N_1 , then the rate of emission involves the spontaneous rate (single photon emission) multiplied by $(N(\nu) + 1)$ and the transition rate, denoted by A_{10} . The decrease is due to stimulated absorption, proportional to the number of absorbing atoms, N_0 , A_{10} , and $N(\nu)$. There is a further loss of photons due to leakage, which is proportional to the number of photons $N(\nu)$. The leakage losses are described by the last term in the equation below, and the cavity must be designed to make τ_0 as large as possible. We write

$$\begin{aligned}\frac{dN(\nu)}{dt} &= N_1(N(\nu) + 1)A_{10} - N_0N(\nu)A_{10} - \frac{N(\nu)}{\tau_0} \\ &= N(\nu)\left\{(N_1 - N_0)A_{10} - \frac{1}{\tau_0}\right\} + N_1A_{10}\end{aligned}\quad (18-2)$$

(in writing the above we have neglected the fact that different levels may have different degeneracies—for the purpose of our discussion this is not relevant) We see from this equation that the number of photons will decrease with time unless

$$N_1 - N_0 > 0 \quad (18-3)$$

In thermal equilibrium we have

$$\frac{N_1}{N_0} = \frac{e^{-E_1/kT}}{e^{-E_0/kT}} = e^{-(E_1 - E_0)/kT} = e^{-h\nu/kT} \quad (18-4)$$

which is always smaller than 1. Thus a laser must operate in a nonequilibrium mode, and we must create an excess population of atoms in the upper level—that is, a *population inversion*. One way of doing this is described below.

Optical Pumping

A way of setting up a population inversion is to use a material in which the process involves transitions between three levels. Figure 18-1 shows a three-level system. The levels have energies $E_2 > E_1 > E_0$. The goal is to create a population inversion in level 1 relative to level 0. The state with energy E_2 must have the property that it is possible to increase its occupation number by some mechanism, and it must decay readily to state E_1 but not to E_0 . The *pumping* of atoms into this excited state is accomplished by a variety of

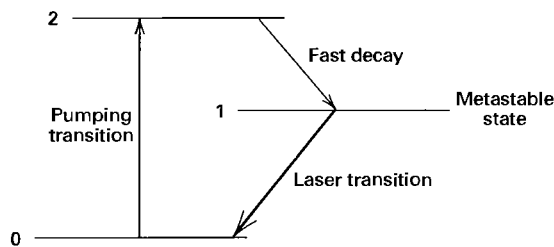


Figure 18-1 Schematic picture of pumping transition, followed by fast decay to a metastable state, from which laser transition occurs.

means, one of which is by subjecting the material to a powerful source of incoherent light, as is done in a *ruby laser*. In that material, the level E_1 is *metastable*: Selection rules slow down spontaneous transitions to the ground level, so that there is time for a large build-up of level 1 atoms by the decay of level 2 atoms.² There are of course many types of lasers and many pumping mechanisms. In this way one may create lasers in different parts of the electromagnetic spectrum. By selecting materials in which the laser transition terminates on a set of closely spaced vibrational molecular levels, one may actually construct *tunable* lasers.

The Cavity

The production of a collimated, narrow beam of photons requires a cylindrical structure for the cavity, with only slightly transparent mirrors at the ends. The mirrors are there to keep the photons inside the cavity to build up the number of photons $n(\nu_{10})$. The losses depend on the value of τ_0 .

We can estimate τ_0 , the “lifetime of photons in the cavity,” as follows: Assume that the radiation only travels back and forth in the cylindrical cavity of length L . If n^* is the refractive index of the medium, then the traversal time is n^*L/c . If the reflection coefficient of the mirror is r (≈ 0.99), then after k traversals, the intensity of the light is reduced to r^k , so that, with $r = 1 - \epsilon$,

$$I_k/I_0 = (1 - \epsilon)^k \approx e^{-k\epsilon}$$

The intensity is reduced to $1/e$ of its original value after $k = 1/\epsilon = 1/(1 - r)$ traversals. Thus the lifetime of the radiation in the cavity is

$$\tau_0 \approx kn^*L/c = n^*L/c(1 - r)$$

The reflection coefficient can be brought closer to unity if the cavity is terminated with Brewster angle windows (which almost perfectly contain one state of polarization of the radiation), beyond which identical spherical mirrors, separated by a distance equal to their common radius of curvature, are placed (Fig. 18-2). Note that the linewidth of the radiation is

$$\Delta\nu = 1/\tau_0 = c(1 - r)n^*L$$

which is very small compared with $c/2n^*L$, the frequency separation of two adjacent modes in a cavity of dimension L . Thus a nearly monochromatic beam is produced.

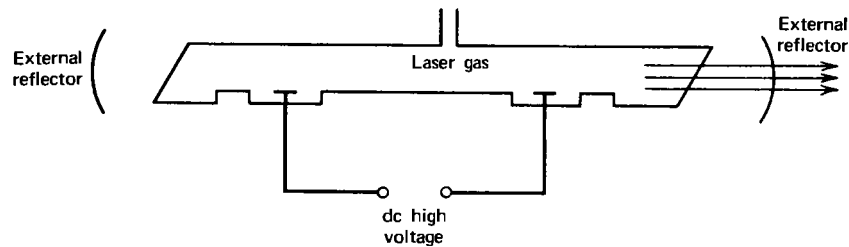


Figure 18-2 Schematic sketch of a gas laser.

²The reader may find more details in *Modern Physics* by J. Bernstein, P. M. Fishbane, and S. Gasiorowicz, Prentice Hall, Englewood Cliffs, N.J., 2000.

18-2 THE COOLING OF ATOMS

In this section we describe one application of lasers that is directly related to the study of atoms—namely, their use to slow down atoms and in this way to reduce the *Doppler broadening* of spectral lines.

Spectral lines can be broadened in several ways; atoms are generally not observed in isolation. In a gas of atoms, there will generally be collisions, and the time between collisions τ_c , if shorter than the mean life of the state under consideration, determines the width of the spectral line, because τ_c is effectively the lifetime of the state, and $\hbar/\tau_c$ is then larger than the natural width of the line. Collision broadening can be reduced by reducing the density (equivalently, the pressure) of the gas of atoms under consideration. There is also *Doppler broadening*. When an atom is moving with speed v , then the frequency of the radiation emitted by the atom is shifted by $\Delta\omega = (v/c)\omega$. If for v we take $v_{\text{rms}} = \sqrt{3kT/M_{\text{atom}}}$, the root-mean-square velocity of atoms in a gas at temperature T , then for hydrogen, for example, $\Delta\omega/\omega \approx 0.3 \times 10^{-6} \sqrt{T}$. This is much larger than the natural value of $\Delta\omega/\omega$, which for the $2p \rightarrow 1s$ line is $\approx 3 \times 10^{-8}$. Thus to overcome the Doppler broadening, the atoms need to be *cooled*. This is done by placing the atom in a laser beam. Consider an atom moving in the positive z -direction with a velocity v . Let the energy difference between the two levels for which a transition occurs be $\hbar\omega_0$, and let the frequency of the (monochromatic) laser beam be ω . We shall take ω a little different from ω_0 , and take the *detuning parameter* $\delta = \omega - \omega_0 < 0$. If the laser beam propagates along the z -axis in the positive z -direction, then the atom “sees” a red-shifted beam, since the origin of the beam appears to be moving away. Thus, the frequency as seen by the atom is $\omega(1 - v/c)$. This means that the amount of detuning of the laser beam away from the resonant frequency is $\omega(1 - v/c) - \omega_0 = \delta - \omega v/c$. For $\delta < 0$, the magnitude of this is larger than the magnitude of δ , and the probability of absorption of the photons is decreased, since the frequency of the light absorbed is farther out on the resonance curve. On the other hand, with another laser beam propagating in the negative z -direction, the detuning is $\omega(1 + v/c) - \omega_0 = \delta + \omega v/c$, so that the frequency of the light absorbed is closer to the peak of the resonance (with $\omega v/c = |\delta|$ one would be at resonance). Thus the probability of absorbing a photon is larger. This means that the atom experiences a net force in the negative z -direction. The atom re-emits the photon, because it decays, but the emission has no preferred direction associated with it, and on the average, it is spherically symmetric.³ Thus there is a net loss of momentum for the atom along the z -direction. If there are three pairs of lasers, aligned along the x -, y -, and z -axes and pairwise of equal intensity, then all of the degrees of freedom are covered.

To make this more quantitative, we calculate the radiation pressure force exerted on the atom by the beam. We assume that there are only two levels that are relevant in the atom. This is a good approximation when the frequency of the oscillating electric field is close to the frequency associated with the energy of excitation of the state above the ground state. The rate at which momentum of magnitude $\hbar\omega/c$ is absorbed was calculated to be

$$\begin{aligned}\Gamma &= \frac{2\pi}{\hbar} |\langle 1 | e\mathbf{r} \cdot \mathbf{E} | 0 \rangle|^2 \delta(E_1 - E_0 - \hbar\omega) \\ &= \frac{2\pi}{\hbar^2} |\langle 1 | e\mathbf{r} \cdot \mathbf{E} | 0 \rangle|^2 \delta(\omega_0 - \omega)\end{aligned}\tag{18-5}$$

³We assume that the intensity of the laser beam is not very large, so that the excitation and the subsequent decay are events that are well separated. Later in this chapter we shall see situations in which the system oscillates rapidly between the ground state and the excited state. In that case, the emission is coherent with the excitation, and directional information is not lost.

where the dipole moment operator $e\mathbf{r}$ gives rise to the transition. We now take into account that the delta function needs to be modified because of the finite linewidth of the excited state, so that

$$\pi\delta(\omega_0 - \omega) \rightarrow \frac{\Gamma/2}{(\omega_0 - \omega)^2 + (\Gamma/2)^2} \quad (18-6)$$

where Γ is the spontaneous decay rate. This means that the radiation pressure is

$$F = \frac{\Delta p}{\Delta t} = \frac{\hbar\omega}{c} \times (\text{rate}) = \frac{\hbar\omega}{c} \frac{2}{\hbar^2} |\langle 1 | e\mathbf{r} \cdot \mathbf{E} | 0 \rangle|^2 \frac{\Gamma/2}{(\omega_0 - \omega)^2 + (\Gamma/2)^2} \quad (18-7)$$

We introduce the dimensionless quantity

$$I = \frac{|\langle 1 | e\mathbf{r} \cdot \mathbf{E} | 0 \rangle|^2}{(\hbar\Gamma/2)^2} \quad (18-8)$$

in terms of which

$$F = \frac{\hbar\omega}{c} I \Gamma \frac{(\Gamma/2)^2}{(\omega_0 - \omega)^2 + (\Gamma/2)^2} \quad (18-9)$$

For a weak field—that is, for I small—the force is very small. For a very large field—that is, for large I —the induced decays overwhelm the spontaneous ones, and under those circumstances the atom oscillates rapidly between the ground state and the excited state.⁴ In particular, the emitted photons are coherent with the absorbed photons, so that the atom does not absorb any net momentum. It turns out that the optimum laser intensity is such that $I \approx 1$.

For an atom moving in the same direction as the beam, there is the additional Doppler shift detuning, so that

$$F = \frac{\hbar\omega}{c} I \Gamma \frac{(\Gamma/2)^2}{(\omega_0 - \omega - \omega v/c)^2 + (\Gamma/2)^2} \quad (18-10)$$

If now a standing wave beam is considered, or, equivalently we add a laser beam with the same frequency ω traveling in the negative z -direction, we get a force in the opposite direction, with the frequency blue-shifted to $\omega(1 + v/c)$. Thus the net force on the beam is

$$F_{\text{net}} = \frac{\hbar\omega}{c} I \Gamma \left(\frac{(\Gamma/2)^2}{(\omega_0 - \omega - \omega v/c)^2 + (\Gamma/2)^2} - \frac{(\Gamma/2)^2}{(\omega_0 - \omega + \omega v/c)^2 + (\Gamma/2)^2} \right)$$

To the lowest order in v/c , one gets

$$F_{\text{net}} = \frac{\hbar\omega I \Gamma}{c} \frac{(\Gamma/2)^2}{\delta^2 + (\Gamma/2)^2} \frac{4\omega\delta}{\delta^2 + (\Gamma/2)^2} v \quad (18-11)$$

Since $\delta < 0$ —that is, the laser frequency is chosen to be a little below the resonance peak ω_0 —the force is in a direction opposite to that of v ; that is, it is a *frictional force*, of the form $F_{\text{net}} = -\beta v$. The force depends on the amount of detuning, and it is largest when $dF/d\delta = 0$ —that is, when $|\delta| = \omega_0 - \omega = \Gamma/2\sqrt{3}$. The atomic motion in all three dimensions must be slowed down, and for this purpose three pairs of lasers are used, to create an environment commonly called *optical molasses*.

The maximum frictional force for $I \approx 1$ is therefore

$$F \approx \sqrt{\frac{27}{4}} \frac{\hbar\omega_0^2}{c^2} v$$

⁴See Section 18-3.

The atoms are also subject to a random force due to the random encounters with the photons that make up the laser beam. They thus behave like particles in the fluid undergoing Brownian motion. The details of the cooling process are beyond the scope of our discussion, but the prediction of the semiclassical argument just presented is that the atoms will be cooled down to a temperature given by

$$T = \frac{\hbar\omega}{k}$$

where k is Boltzmann's constant. This is generally in the range of 2×10^{-4} K. A deeper examination of the details of the process, which takes into account the degeneracy of the excited levels and the polarization of the photons, shows that one should expect to be able to cool atoms to a temperature of order 4×10^{-5} K, which agrees with what has been found experimentally.⁵ In fact, recent improvements in the technology of laser cooling of atoms have yielded atomic temperatures of 4×10^{-7} K. When atoms are cooled to such low temperatures, measurements of the natural linewidth become possible, and this has had a large impact on the testing of quantum electrodynamics predictions.

18-3 THE TWO-LEVEL SYSTEM

Consider an atom in an external time-dependent electric field, which induces transitions between two levels that we label $|a\rangle$ and $|b\rangle$. Their energies can be expressed in terms of angular frequencies ω_a and ω_b so that $E_a = \hbar\omega_a$, $E_b = \hbar\omega_b$ with $E_a > E_b$.

The operator H_0 has the property that

$$\begin{aligned} H_0|a\rangle &= E_a|a\rangle \\ H_0|b\rangle &= E_b|b\rangle \end{aligned} \quad (18-12)$$

Since $|a\rangle$ and $|b\rangle$ form the complete set of states, we may use them to give a matrix representation of H_0 . The matrix is diagonal in the basis that uses the energy eigenstates, so that

$$H_0 = \begin{pmatrix} \hbar\omega_a & 0 \\ 0 & \hbar\omega_b \end{pmatrix} \quad (18-13)$$

The system is perturbed by an external oscillating electric field. We write

$$H_1 = eE_0x \cos \omega t \quad (18-14)$$

with the choice of polarization such that E points in the x -direction. Because of parity conservation, $\langle a|x|a\rangle = \langle b|x|b\rangle = 0$, so that H_1 only has off-diagonal elements. Let us introduce the notation

$$W = eE_0\langle a|x|b\rangle \quad (18-15)$$

Note that W has the dimensions of an energy. This leads to

$$H_1 = \begin{pmatrix} 0 & W \cos \omega t \\ W^* \cos \omega t & 0 \end{pmatrix} \quad (18-16)$$

We shall use the *interaction picture* described in Supplement 15-B [www.wiley.com/college/gasiorowicz]. Briefly, the procedure is to solve the equation satisfied by

$$|\psi_I(t)\rangle = e^{iH_0t/\hbar} |\psi(t)\rangle \quad (18-17)$$

⁵This is discussed in nontechnical detail by C. N. Cohen-Tannoudji and W. D. Phillips in *Phys. Today*, 43(10), 33 (1990). This article also contains many references.

The equation has the form

$$\frac{d}{dt} |\psi_I(t)\rangle = -\frac{i}{\hbar} V(t) |\psi_I(t)\rangle \quad (18-18)$$

with

$$V(t) = e^{iH_0 t/\hbar} H_1 e^{-iH_0 t/\hbar} \quad (18-19)$$

To obtain $V(t)$ in (18-19) we need to calculate $e^{iH_0 t/\hbar}$. With H_0 diagonal, this is very simple and we have

$$e^{iH_0 t/\hbar} = \begin{pmatrix} e^{i\omega_a t} & 0 \\ 0 & e^{i\omega_b t} \end{pmatrix}$$

The calculation of $V(t) = e^{iH_0 t/\hbar} H_1 e^{-iH_0 t/\hbar}$ is just a matter of multiplying three 2×2 matrices. The result is

$$V(t) = \begin{pmatrix} 0 & W e^{i(\omega_a - \omega_b)t} \cos \omega t \\ W^* e^{-i(\omega_a - \omega_b)t} \cos \omega t & 0 \end{pmatrix} \quad (18-20)$$

The expression can be simplified somewhat. We can write

$$e^{i(\omega_a - \omega_b)t} \cos \omega t = \frac{1}{2} (e^{i(\omega_a - \omega_b + \omega)t} + e^{i(\omega_a - \omega_b - \omega)t})$$

Experimentally the ω is *tuned* to the energy difference between the two levels. This means that $\omega_a - \omega_b - \omega \equiv \Delta$ is small. The first term above oscillates very rapidly, and contributes nothing when averaged over time. As in our discussion of *paramagnetic resonance* we drop these terms. Applying this approximation, called the *rotating wave approximation*, we get

$$V(t) = \begin{pmatrix} 0 & (W/2) e^{i\Delta t} \\ (W^*/2) e^{-i\Delta t} & 0 \end{pmatrix} \quad (18-21)$$

Let us now write out (18-18) in detail. With $|\psi_I(t)\rangle$ represented by the column vector

$$|\psi_I(t)\rangle \rightarrow \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \quad (18-22)$$

our equation reads

$$\begin{aligned} i\hbar \frac{da(t)}{dt} &= \frac{W}{2} e^{i\Delta t} b(t) \\ i\hbar \frac{db(t)}{dt} &= \frac{W^*}{2} e^{-i\Delta t} a(t) \end{aligned} \quad (18-23)$$

We differentiate the first of these, and using the second equation, we get, after some minor rearrangements, the equation

$$\frac{d^2 a(t)}{dt^2} - i\Delta \frac{da(t)}{dt} + \frac{|W|^2}{4\hbar^2} a(t) = 0$$

If we substitute $a(t) = e^{i\Omega t}$ into the above, we find that

$$\Omega^2 - \Delta\Omega - \frac{|W|^2}{4\hbar^2} = 0$$

The roots of this are

$$\Omega = \frac{\Delta}{2} \pm \sqrt{\frac{\Delta^2}{4} + \frac{|W|^2}{4\hbar^2}} \equiv \omega_{\pm} \quad (18-24)$$

We now get our solutions as follows: Choose

$$a(t) = a_+ e^{i\omega_+ t} + a_- e^{i\omega_- t} \quad (18-25)$$

Then the equation for $b(t)$ is obtained by differentiating $a(t)$. A little algebra leads to the result

$$b(t) = \frac{2\hbar\omega_+}{|W|} a_+ e^{-i\omega_- t} - \frac{2\hbar\omega_-}{|W|} a_- e^{-i\omega_+ t} \quad (18-26)$$

Let us suppose that the system is initially in the *upper* state, so that $a(0) = 1$ and $b(0) = 0$. This is satisfied if

$$\begin{aligned} a_+ + a_- &= 1 \\ \omega_+ a_+ + \omega_- a_- &= 0 \end{aligned}$$

or, equivalently

$$\begin{aligned} a_+ &= \frac{\omega_-}{\omega_- - \omega_+} = -\frac{\omega_-}{2R} \\ a_- &= -\frac{\omega_+}{\omega_- - \omega_+} = \frac{\omega_+}{2R} \end{aligned} \quad (18-27)$$

where

$$R = \sqrt{\frac{\Delta^2}{4} + \frac{|W|^2}{4\hbar^2}} \quad (18-28)$$

This results in

$$b(t) = \frac{2\hbar\omega_+\omega_-}{a2|W|R} (e^{-i\omega_- t} - e^{-i\omega_+ t}) \quad (18-29)$$

We may use this to calculate the probability that at a time t the system is in the lower state.

$$\begin{aligned} P(t) &= |b(t)|^2 = \frac{4\hbar^2(\omega_+\omega_-)^2}{4|W|^2 R^2} |e^{-i\omega_- t} - e^{-i\omega_+ t}|^2 \\ &= \frac{\hbar^2 \left(\frac{|W|^2}{4\hbar^2} \right)^2}{|W|^2 \left(\frac{\Delta^2}{4} + \frac{|W|^2}{4\hbar^2} \right)} 4 \sin^2 \sqrt{\frac{\Delta^2}{4} + \frac{|W|^2}{4\hbar^2}} t \\ &= \frac{|W|^2}{|W|^2 + \hbar^2 \Delta^2} \sin^2 \sqrt{\frac{\Delta^2}{4} + \frac{|W|^2}{4\hbar^2}} t \end{aligned} \quad (18-30)$$

With perfect *tuning*, the frequency of the external field matches the energy difference between the levels, so that then $\Delta = 0$. In that case

$$P(t) = \sin^2 \frac{|W|}{2\hbar} t \equiv \sin^2 rt \quad (18-31)$$

The system returns to its original state after a time $T = 2\pi/r$. Thus the angular frequency for the *flopping* is given by $2\pi/T = r$. This frequency is known as the *Rabi frequency*.

In our discussion we have treated the external electric field classically. It appeared in the form

$$W = eE_0\langle b|x|a\rangle$$

If we treat the electromagnetic field quantum mechanically, then $\mathbf{E}_0$ is an operator, and the above quantity is to be replaced by

$$W_N = e\langle N+1|\mathbf{E}_0|N\rangle\langle b|x|a\rangle = \sqrt{N+1}W \quad (18-32)$$

as discussed in Supplement 18-A [www.wiley.com/college/gasiorowicz]. In the above we have explicitly assumed that we have N photons in the upper state, and $N+1$ photons in the lower state. This changes the *Rabi frequency* by a factor of $\sqrt{N+1}$ to $r\sqrt{N+1}$. If the initial state has $N+1$ photons and the final state N photons, then it is the *lowering* operator that plays a role, and the frequency is $r\sqrt{N}$.

The study of single atoms in a cavity in which there is only one mode of an oscillating electric field present has been made possible through the invention and construction of radio-frequency traps by H. Dehmelt and collaborators. The oscillations predicted above have been experimentally established.

18-4 THE THREE-LEVEL SYSTEM

The three-level system (Fig. 18-3) has many interesting features. The three levels, A , B , C with energies $\hbar\omega_a > \hbar\omega_b > \hbar\omega_c$ is placed in a pair of electric fields. One of them is characterized by a frequency ω_1 close to the difference $\omega_a - \omega_b$, the other by a frequency ω_2 close to the difference $\omega_a - \omega_c$. The perturbing Hamiltonian is thus

$$H_1 = eE_1x \cos \omega_1 t + eE_2x \cos \omega_2 t \quad (18-33)$$

If the state vector (in the interaction picture) is represented by

$$|\psi_1(t)\rangle \rightarrow \begin{pmatrix} a(t) \\ b(t) \\ c(t) \end{pmatrix} \quad (18-34)$$

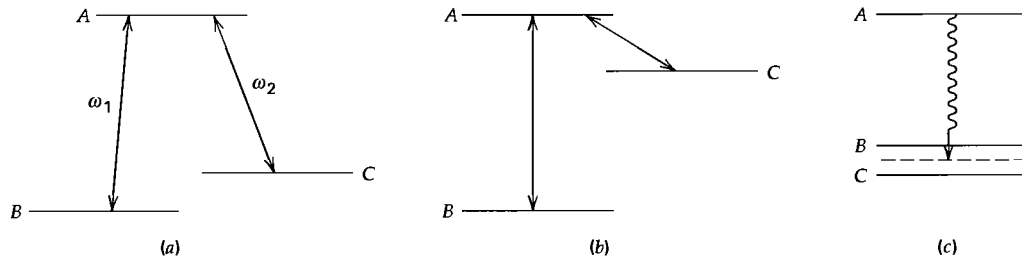


Figure 18-3 (a) Three-level system, driven by lasers with angular frequencies ω_1 and ω_2 creating a *dark state* $|A\rangle$ by a coherent superposition of the two lower states. (b) Three-level system arranged so that the probability of exciting an atom in the $|B\rangle$ state to the $|A\rangle$ state is very small, leading to transparency to photons of frequency ω_1 . (c) Phase relation between states $|B\rangle$ and $|C\rangle$ is so arranged that the amplitudes for the transitions to $|A\rangle$ cancel.

then our task is to calculate the amplitudes for the occupation of these three states, $a(t)$, . . . The details of the calculation in the simplifying limit of perfect tuning, so that

$$\begin{aligned}\omega_1 &= \omega_a - \omega_b \\ \omega_2 &= \omega_a - \omega_c\end{aligned}\tag{18-35}$$

are relegated to Supplement 18-B [www.wiley.com/college/gasiorowicz]. We find that different initial conditions allow for interesting phenomena. The calculation shows that:

1. If the states $|b\rangle$ and $|c\rangle$ lie together, and the field strengths are chosen such that

$$eE_1\langle b|x|a\rangle = eE_2\langle c|x|a\rangle\tag{18-36}$$

then, with $a(0) = 0$, $b(0) = -c(0) = 1/\sqrt{2}$ we find that $a(t) = 0$ for *all times*. This means that the state A is never excited. It is therefore called a *dark state*.

2. If we consider a situation in which $E_1 \gg E_2$, one can show that with the initial conditions $a(0) = b(0) = 0$, the state A is very unlikely to be excited by a $C \rightarrow A$ induced transition. This means that for light of the frequency $\omega_a - \omega_c$ the probability of a photon absorption is very small, and this in turn means that we get *electromagnetically induced transparency* at that frequency.
3. We do not discuss this in Supplement 18-B [www.wiley.com/college/gasiorowicz], but it is possible to arrange the phases of the external fields such that we can get *lasing* without creating a population inversion.

Basically the idea is that the probability of transitions from A to B and C is the sum of the probabilities $P(A \rightarrow B)$ and $P(A \rightarrow C)$. On the other hand, the probability of transitions that populate A is proportional to $|b(t) + c(t)|^2$ and it is possible to arrange that the two amplitudes interfere. This means that photons are *not* absorbed by the states B and C (Fig. 18-3).⁶

18-5 OBSERVATION OF QUANTUM JUMPS

The invention of the traps in which single ions could be studied allows for a variety of new techniques to study atoms. One ingenious idea was first proposed by H. Dehmelt and observed by several experimental groups in the past decade. The principle of the experiment is the following: Consider a three-level system, consisting of the ground state $|0\rangle$ and two excited states $|1\rangle$ and $|2\rangle$. Transitions between the states $|0\rangle$ and $|1\rangle$ are allowed, while those between $|2\rangle$ and the ground state are forbidden (but not absolutely forbidden), so that the state $|2\rangle$ is metastable. A strong light source (laser) tuned to the angular frequency $\omega_{10} = (E_1 - E_0)/\hbar$ is now directed at the atom, as is a weak light source tuned to the frequency $\omega_{20} = (E_2 - E_0)/\hbar$. The number of transitions between the ground state and the allowed excited state is very large. Effectively the strong laser field excites the electron to the state $|1\rangle$ at a very rapid rate, and the electron decays to the ground state also at a very rapid rate. Thus a continuous signal of light emitted by the atom is observed. This is just a manifestation of the Rabi oscillations discussed in Section 18-4. Very occasionally the electron is excited by the weak laser to the state $|2\rangle$. Since the electron now finds itself in a metastable state, the time until it decays to the ground state may be as long as seconds. *During that time there is no fluorescence*—that is, the atom is dark. When the

⁶Much of this material is described in more detail in the excellent book *Quantum Optics* by M. O. Scully and M. S. Zubairy, Cambridge University Press (1997).

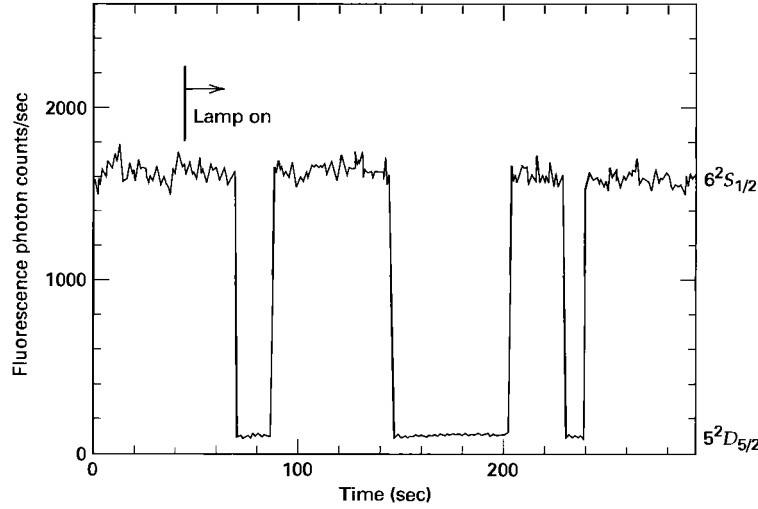


Figure 18-4 A typical trace of fluorescence from showing the quantum jumps. The atom is definitely known to be in the shelf level during the low fluorescence periods. (From W. Nagourney, J. Sandberg, and H. Dehmelt, *Phys. Rev. Lett.* **56**, 2797 (1986), by permission.)

electron finally does decay to the ground state, it is immediately excited by the strong laser field to the allowed excited state, and rapidly decays, giving rise to a resumption of the fluorescent radiation. Effectively the fluorescence monitors the quantum jumps between the ground state and the metastable state (Fig. 18-4).

Let A_{10} and A_{20} be the spontaneous transition rates for $|1\rangle \rightarrow |0\rangle$ and $|2\rangle \rightarrow |0\rangle$, respectively. The conditions on the transitions imply that $A_{10} \gg A_{20}$. If the energy densities of the laser beams are U_1 and U_2 , respectively, then the rate equations for the probability that the atom is in the excited state $|1\rangle$ (assuming no degeneracies so that all the g_i are 1), is

$$\frac{dP_1}{dt} = -P_1 A_{10}(1 + N_1 U_1) + P_0 A_{10} N_1 U_1 \quad (18-37)$$

Here N_1 is the number of photons with frequency ω_{10} present. The equation describes the loss of probability due to spontaneous and stimulated emission, and its buildup due to stimulated absorption from the ground state, the latter proportional to the probability P_0 that the electron is in the ground state. Similarly the rate equation for the probability that the atom is in the metastable state $|2\rangle$.

$$\frac{dP_2}{dt} = -P_2 A_{20}(1 + N_2 U_2) + P_0 A_{20} N_2 U_2 \quad (18-38)$$

where N_2 is the number of photons with frequency ω_{20} that are present.

The probabilities add up to 1, so that $P_0 + P_1 + P_2 = 1$. If the laser coupling the ground state to the excited state $|1\rangle$ is intense, so that U_1 is very large, then equation (18-37) implies that $P_1 = P_0$ after a very short time. Under these circumstances, we may ask for the probability that the metastable state is excited. If we call this $P_+ (= P_2)$, and write the probability that it is *not* excited as $P_- (= 1 - P_+ = P_0 + P_1 = 2P_0)$, we have

$$\frac{dP_+}{dt} = -\Gamma_- P_+ + \Gamma_+ P_- \quad (18-39)$$

with $\Gamma_+ = \frac{1}{2} A_{20} N_2 U_2$ and $\Gamma_- = A_{20}(1 + N_2 U_2)$. Since $P_+ + P_- = 1$, the equation for P_- follows automatically. We may view these equations as representing a two-level system in which the upward transition rate is Γ_+ and the downward transition rate is Γ_- .

Experimentally interesting quantities are the probabilities that in the time interval from t to $t + T$ no transitions have occurred, and that at the end of the interval the electron ends either in the excited state, P_{0+} or in the ground state P_{0-} . A little reflection shows that once the experiment starts and the laser beam intensities are independent of time, these probabilities only depend on the length of the time interval T . We can write rate equations for these probabilities, and these are

$$\frac{dP_{0+}}{dT} = -\Gamma_- P_{0+} \quad (18-40)$$

$$\frac{dP_{0-}}{dT} = -\Gamma_+ P_{0-} \quad (18-41)$$

The “initial conditions” for these equations require that we know $P_{0\pm}(T = 0)$. Suppose that the intensity of fluorescence turns off at some time t . We choose that time to denote $T = 0$, so that $P_{0+}(T = 0) = 1$. The solution to the equation with this initial condition is

$$P_{0+}(T) = e^{-\Gamma_- T} \quad (18-42)$$

This is the probability that the signal is still “off” after a time T . Similarly one can show that if the fluorescence has just turned on, then the probability that it is still on after a time interval T is

$$P_{0-}(T) = e^{-\Gamma_+ T} \quad (18-43)$$

A statistical analysis of the distribution of on and off times as shown in Fig. 18-5 can be used to measure A_2 . For very long lived states, *direct* measurements of A_2 are very difficult, since the rate at which photons are emitted is very small, and they can be emitted in all directions, so counting them is a very slow process.

It is interesting to note that quantum mechanics usually involves the study of a large ensemble of identical systems. In this case, one is studying a single atom and the ensemble is replaced by rebuilding over and over again that single member of the ensemble with identical initial conditions. This is generally not possible when radiation

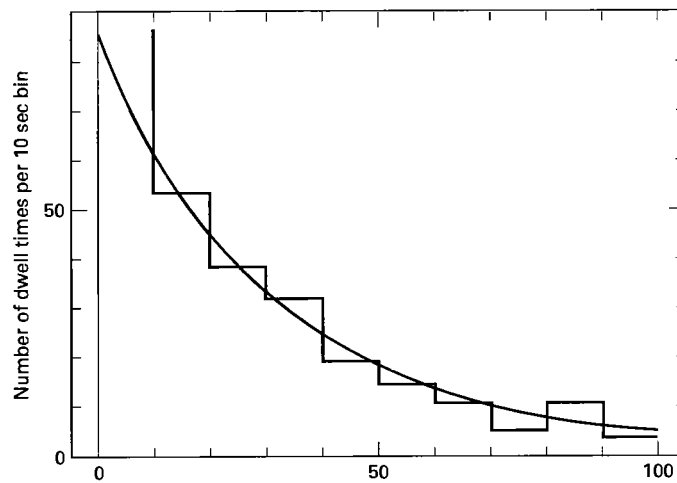


Figure 18-5 Dwell time in shelf level (s). (From W. Nagourney, J. Sandberg, and H. Dehmelt, *Phys. Rev. Lett.* **56**, 2797 (1986), by permission.)

is emitted into a continuum of states, but it is possible in the special case of a single mode electric field.

8-6 THE MÖSSBAUER EFFECT

An atom (or any other quantum system) can act as a very accurate clock, since its transitions are signaled by radiation of a very well-defined frequency. If the only limitation were the natural linewidth, great accuracy could be obtained.

Unfortunately, as stated in our discussion of the cooling of atoms, the motion of atoms leads to the broadening of lines by the Doppler effect. One might think that the use of a liquid or solid source would eliminate this effect, but then broadening caused by the effect of neighboring atoms is just as harmful. Let us consider nuclear transitions. A nucleus such as ^{191}Ir emits a γ -ray of energy of the order of 100 keV, with a lifetime of 10^{-10} s. This corresponds to

$$\frac{\Delta\omega}{\omega} = \frac{\Delta E}{E} = \frac{\hbar/\tau}{E} \approx \frac{10^{-34}/10^{-10}}{10^5 \times 1.6 \times 10^{-19}} \approx 0.6 \times 10^{-10} \quad (18-44)$$

There will, unfortunately, be a recoil shift of the line. The γ -ray carries off momentum $\hbar\omega/c$, and the nucleus, to conserve momentum, must recoil with the same momentum. This gives rise to a recoil energy

$$\Delta E = \frac{P_{\text{recoil}}^2}{2M} = \frac{1}{2M} \left(\frac{\hbar\omega}{c} \right)^2 \quad (18-45)$$

and thus a decrease in the energy radiated. The fractional change in frequency is

$$\frac{\Delta E}{\hbar\omega} \approx \frac{\hbar\omega}{2Mc^2} \approx \frac{10^{-1}(\text{MeV})}{2 \times 940 \times 191(\text{MeV})} \approx 3 \times 10^{-7} \quad (18-46)$$

The observation of radiation of this energy cannot be carried out with the conventional, extremely accurate spectroscopic methods, but must utilize a detector that is extremely “well tuned” to the radiation. This is best done by using the same material (e.g., ^{191}Ir) as an absorber. The absorption will be very much enhanced at the “resonant” frequency at which the radiation is emitted, but here, too, there will be a recoil shift. The overall shift is thus $\Delta\omega/\omega \approx 6 \times 10^{-7}$. Thus, the “fine tuning” does not work, since the line is shifted by far more than the width, which is of the order of $10^{-10} \omega$. One could try to compensate for the recoil by moving the emitter with the recoil velocity. This is given by

$$\frac{v}{c} = \frac{P_{\text{recoil}}}{Mc} = \frac{\hbar\omega/c}{Mc} = 2 \frac{\hbar\omega}{2Mc^2} \approx 6 \times 10^{-7} \quad (18-47)$$

that is, $v = 1.8 \times 10^2$ m/s. This presents technical difficulties, but it has been achieved with an ultracentrifuge.

A major breakthrough came with the discovery by Mössbauer in 1958 that under certain conditions there is a high probability of *recoilless emission*. The emission is not recoilless, of course; however, the recoil is not taken up by the nucleus, but instead by a large part of the crystal that the nucleus is imbedded in. Since the mass of the nucleus is 10^{22} times smaller than that of the crystal, the recoil energy is completely negligible. To get some intuition about what is happening, let us consider the nucleus as moving in a

harmonic oscillator well, with characteristic frequency ω_0 . The energy levels of the oscillator are

$$E_n = \hbar\omega_0 (n_x + n_y + n_z + \frac{3}{2}) \quad (18-48)$$

The harmonic well is just an approximate description of the crystalline forces that are responsible for the properties of the lattice. If the forces that tie the nucleus to its neighbors are strong—if the “springs” are stiff—then ω_0 is large; if the “springs” are soft, then ω_0 is small. In terms of level spacing, a “stiff spring” has widely separated levels—that is, a low density of states—whereas a “soft spring” has a high density of states. Let us now consider the matrix element for a transition from a nuclear state described by $\Psi_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ to a nuclear state described by $\Psi_f(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$, and we take the interaction to be

$$-\frac{e}{M} \sum_{\text{protons}} \mathbf{p}_k \cdot \mathbf{A}_k(\mathbf{r}_k, t) \quad (18-49)$$

The matrix element then is proportional to

$$-\frac{e}{M} \int \dots \int d^3\mathbf{r}_1, \dots, d^3\mathbf{r}_N \Psi_f^*(\mathbf{r}_1, \dots, \mathbf{r}_N) \sum_k \boldsymbol{\varepsilon} \cdot \mathbf{p}_k e^{-i\mathbf{k} \cdot \mathbf{r}_k} \Psi_i(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad (18-50)$$

If we introduce the center-of-mass coordinate $\mathbf{R} = (1/N) \sum_i \mathbf{r}_i$ then (a) the interaction term takes the form

$$-\frac{e}{M} e^{-i\mathbf{k} \cdot \mathbf{R}} \sum_{\text{protons}} \boldsymbol{\varepsilon} \cdot \mathbf{p}_k e^{-i\mathbf{k} \cdot \mathbf{p}_k} \quad (18-51)$$

where $\mathbf{p}_i = \mathbf{r}_i - \mathbf{R}$, and (b) the nuclear wave function decomposes into a product describing the internal motion and the motion of the nuclear center of mass in the harmonic potential

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \psi_{n_x, n_y, n_z}(\mathbf{R}) \phi(\mathbf{p}_1, \dots, \mathbf{p}_{N-1}) \quad (18-52)$$

Thus the matrix element (18-50) becomes

$$\begin{aligned} & -\frac{e}{M} \int d^3\mathbf{R} \psi_{n_f}^*(\mathbf{R}) e^{-i\mathbf{k} \cdot \mathbf{R}} \psi_{n_i}(\mathbf{R}) \\ & \times \int d^3\mathbf{p}_1, \dots, d^3\mathbf{p}_{N-1} \phi_f^*(\mathbf{p}_1, \dots, \mathbf{p}_{N-1}) \sum_{\text{protons}} \boldsymbol{\varepsilon} \cdot \mathbf{p}_k e^{-i\mathbf{k} \cdot \mathbf{p}_k} \phi_i(\mathbf{p}_1, \dots, \mathbf{p}_{N-1}) \end{aligned} \quad (18-53)$$

We may write this in the form

$$M = M_{\text{internal}} \int d^3\mathbf{R} \psi_{n_f}^*(\mathbf{R}) e^{-i\mathbf{k} \cdot \mathbf{R}} \psi_0(\mathbf{R}) \quad (18-54)$$

where we have set $n_i = 0$, since the initial state is in the ground state of the lattice. The probability that the radiative transition leaves the nucleus in the lattice ground

$$\begin{aligned} P_0(k) &= \frac{|\mathbf{M}_{\text{int}}|^2 \left| \int d^3\mathbf{R} \psi_0^*(\mathbf{R}) e^{-i\mathbf{k} \cdot \mathbf{R}} \psi_0(\mathbf{R}) \right|^2}{|\mathbf{M}_{\text{int}}|^2 \sum_{n_f} \left| \int d^3\mathbf{R} \psi_{n_f}^*(\mathbf{R}) e^{-i\mathbf{k} \cdot \mathbf{R}} \psi_0(\mathbf{R}) \right|^2} \\ &= \left| \int d^3\mathbf{R} \psi_0^*(\mathbf{R}) e^{-i\mathbf{k} \cdot \mathbf{R}} \psi_0(\mathbf{R}) \right|^2 \end{aligned} \quad (18-55)$$

In the last step we replaced the sum in the denominator by unity, using completeness.⁷ To calculate this, we use the normalized ground-state wave function of the oscillator. We found in Chapter 6 that the one-dimensional ground-state wave function is

$$\psi_0(x) = \left(\frac{m\omega_0}{\pi\hbar}\right)^{1/4} e^{-m\omega_0 x^2/2\hbar}$$

Hence, for three dimensions we have

$$\psi_0(\mathbf{R}) = \psi_0(x) \psi_0(y) \psi_0(z) = \left(\frac{m\omega_0}{\pi\hbar}\right)^{3/4} e^{-m\omega_0 \mathbf{R}^2/2\hbar} \quad (18-56)$$

We thus calculate

$$\left| \left(\frac{M_N\omega_0}{\pi\hbar}\right)^{3/2} \int d^3\mathbf{R} e^{-M_N\omega_0 \mathbf{R}^2/\hbar} e^{-i\mathbf{k}\cdot\mathbf{R}} \right|^2$$

where M_N is the mass of the nucleus. We get

$$\begin{aligned} P_0 &= \left(\frac{M_N\omega_0}{\pi\hbar}\right)^3 \left| \int d^3\mathbf{R} e^{-M_N\omega_0 \mathbf{R}^2/\hbar} e^{i\mathbf{k}\cdot\mathbf{R}} e^{-k^2\hbar/4M_N\omega_0} \right|^2 \\ &= e^{-\hbar^2 k^2/2M_N\omega_0} \\ &= e^{-\text{recoil energy/level spacing}} \end{aligned} \quad (18-57)$$

since $P_{\text{recoil}} = \hbar k$ and $\hbar\omega_0$ is the level spacing in the lattice. Thus, if the level spacing is large—that is, we have a stiff spring—recoilless emission becomes more probable. The model of the lattice that was used here, that of each nucleus moving in its own harmonic potential, is the Einstein model of a lattice, and the frequency ω_0 is the so-called *Debye frequency*, so that we should really replace ω_0 by ω_D , which is related to the Debye temperature T_D by

$$\hbar\omega_D = kT_D \quad (18-58)$$

A more accurate treatment of the lattice using the Debye model for its description merely changes the exponent by a factor of 3/2.

It is not quite correct to say that the whole crystal recoils; instead, in a time τ equal to the lifetime of the transition (1.4×10^{-7} s for Fe^{57}), only a region of the crystal of magnitude

$$L = v_s \tau$$

where v_s is the velocity of propagation of a lattice disturbance, (i.e., the velocity of sound) absorbs the recoil. Now a reasonable estimate of v_s is given by

$$v_s \simeq \frac{a\omega_D}{2\pi}$$

⁷The formal proof is quickest. We have

$$\sum_{n_f} |\langle n_f | e^{i\mathbf{k}\cdot\mathbf{R}} | 0 \rangle|^2 = \sum_{n_f} \langle 0 | e^{-i\mathbf{k}\cdot\mathbf{R}} | n_f \rangle \langle n_f | e^{i\mathbf{k}\cdot\mathbf{R}} | 0 \rangle$$

Using

$$1 = \sum |n_f\rangle\langle n_f|$$

one gets

$$\langle 0 | e^{-i\mathbf{k}\cdot\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} | 0 \rangle = 1$$

where a is the lattice spacing. Thus

$$\frac{L}{a} \approx \frac{\omega_D \tau}{2\pi}$$

and with $\omega_D \approx 10^{13} \text{ s}^{-1}$, the number of nuclei absorbing the recoil, $\sim (L/a)^3$ is still enormous.

The preceding estimates, combined with the uncertainty relation, may be used to show that it is not possible to determine whether it is a single nucleus that “really” recoils. To measure the recoil energy $\hbar^2 k^2 / 2M_N$ takes a time of the order of

$$\Delta t \gg \frac{\hbar}{(\hbar^2 k^2 / 2M_N)}$$

The condition for the Mössbauer effect to occur is that

$$\frac{\hbar^2 k^2}{2M_N} < \hbar \omega_D$$

Hence

$$\Delta t \gg \frac{1}{\omega_D}$$

During that time the disturbance will have traveled a distance

$$d \approx v_s \Delta t \sim \frac{a \omega_D}{2\pi} \Delta t \gg \frac{a}{2\pi}$$

that is, over a distance covering many nuclei.

The question arises of how did we manage to get away from the problem of recoil and momentum conservation by talking about the energy states of the nucleus in the crystal lattice? Where does it say that the crystal absorbed the momentum? The quantum mechanical answer is that, if we want to talk about momentum, we should work in a momentum representation. This, however, is complicated, since it is difficult to describe the crystal forces in that representation. What one must do is to decompose the crystal motion (the crystal is just a lot of oscillators with nearest neighbor “springs”) into normal modes and quantize these. The quanta of the lattice motion, analogs of photons, are the *phonons*. Recoilless emission then means a transition in which phonons are not emitted. The resulting formula is very similar to (18-57). Under these circumstances, the recoil broadening is infinitesimal compared to the natural linewidth. There is still Doppler broadening because of the thermal motion, but this can be handled by cooling the emitter and absorber.

Recoilless emitters provide us with a superb clock, and research utilizing the Mössbauer effect has been done in many fields, such as solid-state physics and chemistry. We will mention just one application, the terrestrial measurement of the gravitational red shift. It follows from the equivalence principle that a photon will have its frequency shifted by

$$\frac{\Delta \omega}{\omega} = \frac{gx}{c^2} \quad (18-59)$$

if it falls through a height x . This can be compensated by a recoil of velocity v .

$$v^2 = 2gx \quad (18-60)$$

(If the photon and the absorber were to fall freely together there would be resonant absorption.) If the absorber or the source is allowed to oscillate rapidly—one uses a transducer—and the absorption curve is correlated with the oscillations, it is possible to check the gravitational shift. Since the velocity, for a separation $x = 20$ m, is of the order of ~ 20 m/s, the experiment is feasible, and was carried out by several groups. Within the errors, the effect is confirmed. For example, for Fe^{57} the predicted shift is $\Delta\omega/\omega = 4.92 \times 10^{-15}$, and the experimental shift found by Pound and Rebka is $(5.13 \pm 0.51) \times 10^{-15}$. A similar experiment in which the energy shift of the γ -ray emitted by Fe^{57} accelerated on a rapidly rotating turntable was measured again yielded results in agreement with the equivalence principle.

Supplement 18-A

Quantizing the Electromagnetic Field Without Frills

Our goal is to express the electromagnetic field in terms of photons. In terms of the vector potential $\mathbf{A}(\mathbf{r}, t)$ we write

$$\begin{aligned}\mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{A}(\mathbf{r}, t) \\ \mathbf{E}(\mathbf{r}, t) &= -\frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t}\end{aligned}\quad (18A-1)$$

As elsewhere in the book, we work in the gauge $\nabla \cdot \mathbf{A}(\mathbf{r}, t) = 0$; we have also made the choice $\phi(\mathbf{r}) = 0$, possible in the absence of charges.

The energy carried by the electromagnetic field is

$$H = \frac{\varepsilon_0}{2} \int d^3r (\mathbf{E}^2(\mathbf{r}, t) + c^2 \mathbf{B}^2(\mathbf{r}, t)) \quad (18A-2)$$

Let us now expand $\mathbf{A}(\mathbf{r}, t)$ in a Fourier series in a cubical box of volume $V = L^3$. We write

$$\mathbf{A}(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}, \lambda} f(\mathbf{k}) \varepsilon_{\lambda}(\mathbf{k}) [A_{\lambda}^+(\mathbf{k}) e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega t)} + A_{\lambda}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}] \quad (18A-3)$$

This will obey the wave equation provided that

$$\omega^2 = \mathbf{k}^2 c^2 \quad (18A-4)$$

The gauge-fixing condition implies that

$$\mathbf{k} \cdot \varepsilon^{(\lambda)}(\mathbf{k}) = 0 \quad (18A-5)$$

This means that the *polarization vectors* $\varepsilon^{(\lambda)}(\mathbf{k})$ ($\lambda = 1, 2$) are perpendicular to the direction of propagation of the wave $\mathbf{k}$; that is, the polarization is *transverse*.

Let us now calculate the energy in terms of the $A_{\lambda}(\mathbf{k})$ and $A_{\lambda}^+(\mathbf{k})$. We get for the first term involving the electric field,

$$\begin{aligned}\frac{1}{V} \int d^3r \sum_{\mathbf{k}, \lambda} \sum_{\mathbf{q}, \lambda'} f(\mathbf{k}) f(\mathbf{q}) \varepsilon_{\lambda}(\mathbf{k}) \varepsilon_{\lambda'}(\mathbf{q}) & [-i\omega A_{\lambda}^+(\mathbf{k}) e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega t)} + i\omega A_{\lambda}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}] \\ & \times [-i\omega' A_{\lambda'}^+(\mathbf{q}) e^{-i(\mathbf{q} \cdot \mathbf{r} - \omega' t)} + i\omega' A_{\lambda'}(\mathbf{q}) e^{i(\mathbf{q} \cdot \mathbf{r} - \omega' t)}]\end{aligned}\quad (18A-6)$$

We use the notation $\omega' = |\mathbf{q}|c$. The values of $\mathbf{k}$ and $\mathbf{q}$ are determined by the fact that we are expanding in a box, and we choose periodic boundary conditions. Thus in any direction we require that

$$e^{ikx} = e^{ik(x+L)} \quad (18A-7)$$

so that $k_1 L = 2\pi n_1$, and so on. Thus the summations are over integers $n_1, n_2, n_3 = 1, 2, 3, \dots$. We do not need to sum over negative integers, since these are contained in the $\mathbf{A}^+$ terms. Now

$$\begin{aligned} \frac{1}{V} \int d^3r e^{i(\mathbf{k}+\mathbf{q})\mathbf{r}} &= 0 \\ \frac{1}{V} \int d^3r e^{i(\mathbf{k}-\mathbf{q})\mathbf{r}} &= \delta_{\mathbf{k}\mathbf{q}} \end{aligned} \quad (18A-8)$$

We also choose

$$\varepsilon_\lambda(\mathbf{k}) \cdot \varepsilon_{\lambda'}(\mathbf{k}) = \delta_{\lambda\lambda'} \quad (18A-9)$$

so that the two polarization directions corresponding to $\lambda = 1, 2$ are perpendicular to each other. This yields the result

$$H_{(E)} = \frac{\varepsilon_0}{2} \sum_{\mathbf{k}, \lambda} f^2(\mathbf{k}) \omega^2 (A_\lambda(\mathbf{k}) A_\lambda^+(\mathbf{k}) + A_\lambda^+(\mathbf{k}) A_\lambda(\mathbf{k})) \quad (18A-10)$$

The second term, involving $\mathbf{B}$, requires the calculation of

$$\begin{aligned} &\frac{1}{V} \sum_{\mathbf{k}, \lambda} \sum_{\mathbf{q}, \lambda'} f(\mathbf{k}) f(\mathbf{q}) (\mathbf{k} \times \varepsilon_\lambda(\mathbf{k})) \cdot (\mathbf{q} \times \varepsilon_{\lambda'}(\mathbf{q})) \\ &\times \{iA_\lambda^+(\mathbf{k}) e^{-i(\mathbf{k}\mathbf{r}-\omega t)} - iA_\lambda(\mathbf{k}) e^{i(\mathbf{k}\mathbf{r}-\omega t)}\} \{iA_{\lambda'}^+(\mathbf{q}) e^{-i(\mathbf{q}\mathbf{r}-\omega' t)} - iA_{\lambda'}(\mathbf{q}) e^{i(\mathbf{q}\mathbf{r}-\omega' t)}\} \end{aligned} \quad (18A-11)$$

The integration over all the spatial coordinates again yields

$$\begin{aligned} \frac{1}{V} \int d^3r e^{i(\mathbf{k}+\mathbf{q})\mathbf{r}} &= 0 \\ \frac{1}{V} \int d^3r e^{i(\mathbf{k}-\mathbf{q})\mathbf{r}} &= \delta_{\mathbf{k}\mathbf{q}} \end{aligned} \quad (18A-12)$$

The vector identity

$$(\mathbf{k} \times \varepsilon_\lambda(\mathbf{k})) \cdot (\mathbf{k} \times \varepsilon_{\lambda'}(\mathbf{k})) = k^2 \varepsilon_\lambda(\mathbf{k}) \cdot \varepsilon_{\lambda'}(\mathbf{k}) - (\mathbf{k} \cdot \varepsilon_\lambda(\mathbf{k}))(\mathbf{k} \cdot \varepsilon_{\lambda'}(\mathbf{k})) \quad (18A-12)$$

with the help of the transversality condition and $\varepsilon_\lambda(\mathbf{k}) \cdot \varepsilon_{\lambda'}(\mathbf{k}) = \delta_{\lambda\lambda'}$ yields $k^2 \delta_{\lambda\lambda'}$. This means that the part of H involving B , when multiplied by c^2 , yields the same factor as the E^2 term. We thus get

$$H = \varepsilon_0 \sum_{\mathbf{k}, \lambda} f^2(\mathbf{k}) \omega^2 [A_\lambda(\mathbf{k}) A_\lambda^+(\mathbf{k}) + A_\lambda^+(\mathbf{k}) A_\lambda(\mathbf{k})] \quad (18A-13)$$

The form looks very much like the sum of terms in the simple harmonic oscillator. In fact, had we chosen $f(\mathbf{k})$ such that

$$\varepsilon_0 f^2(\mathbf{k}) \omega^2 = \frac{1}{2} \hbar \omega$$

that is,

$$f(\mathbf{k}) = \sqrt{\frac{\hbar}{2\varepsilon_0\omega}}$$

we would have obtained

$$H = \sum_{\mathbf{k}, \lambda} \frac{1}{2} \hbar \omega [A_\lambda(\mathbf{k}) A_\lambda^+(\mathbf{k}) + A_\lambda^+(\mathbf{k}) A_\lambda(\mathbf{k})] \quad (18A-14)$$

Let us now assume that the $A_\lambda(\mathbf{k})$ and $A_\lambda^\dagger(\mathbf{k})$ are *operators* that obey the same commutation relations as the operators A and $A^\dagger$ for the simple harmonic oscillator problem; that is

$$[A_\lambda(\mathbf{k}), A_{\lambda'}^\dagger(\mathbf{q})] = \delta_{\lambda\lambda'} \delta_{\mathbf{k}\mathbf{q}} \quad (18A-15)$$

Then we get

$$H = \sum_{\mathbf{k}, \lambda} \hbar\omega ((A_\lambda^\dagger(\mathbf{k}) A_\lambda(\mathbf{k}) + \frac{1}{2})) \quad (18A-16)$$

Actually the second term $\sum_{\mathbf{k}, \lambda} \hbar\omega/2$ is infinite. We sweep this problem under the rug by observing that all energy measurements are measurements of energy differences. We thus concentrate on

$$H = \sum_{\mathbf{k}, \lambda} \hbar\omega A_\lambda^\dagger(\mathbf{k}) A_\lambda(\mathbf{k}) \quad (18A-17)$$

We may now go through the same steps that we did with the harmonic oscillator. For each value of $\mathbf{k}$ and λ we have creation and annihilation operators, and for each value of $\mathbf{k}$ and λ we have states of zero, one, two, . . . photons. The zero photon state, the *vacuum state*, is $|0\rangle$, defined by

$$A_\lambda(\mathbf{k})|0\rangle = 0 \quad (18A-18)$$

A state with n photons of momentum $\hbar\mathbf{k}$ and energy $\hbar\omega$ is given by

$$\frac{1}{\sqrt{n!}} (A_\lambda^\dagger(\mathbf{k}))^n |0\rangle \quad (18A-19)$$

What we have done is to decompose the electromagnetic field into modes, each of which represents photons. Thus

$$\mathbf{E}(\mathbf{r}, t) = \frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}, \lambda} \sqrt{\frac{\hbar\omega}{2\varepsilon_0}} \varepsilon_\lambda(\mathbf{k}) \{-iA_\lambda^\dagger(\mathbf{k}) e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + iA_\lambda(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\} \quad (18A-20)$$

will annihilate or create a single photon.

As a check we can calculate the momentum carried by the electromagnetic field. We need to calculate

$$\mathbf{P} = \varepsilon_0 \int d^3r (\mathbf{E}(\mathbf{r}, t) \times \mathbf{B}(\mathbf{r}, t)) \quad (18A-21)$$

Using the expressions obtained above, we find, after a page of algebra, that

$$\mathbf{P} = \sum_{\mathbf{k}, \lambda} \hbar\mathbf{k} (A_\lambda^\dagger(\mathbf{k}) A_\lambda(\mathbf{k}) + 1) = \sum_{\mathbf{k}, \lambda} \hbar\mathbf{k} A_\lambda^\dagger(\mathbf{k}) A_\lambda(\mathbf{k}) \quad (18A-22)$$

The last step follows from the fact that $\sum_{\mathbf{k}} \mathbf{k} = 0$ by symmetry. We may thus interpret the product $A_\lambda^\dagger(\mathbf{k}) A_\lambda(\mathbf{k}) \equiv N_\lambda(\mathbf{k})$ as the operator representing the number of photons of momentum k and polarization λ .

Supplement 18-B

Details of the Three-Level System

The three-level system has many interesting features, so that we discuss it in some detail. The system of three levels, A, B, C with energies $\hbar\omega_a > \hbar\omega_b > \hbar\omega_c$ is placed in a set of electric fields. One of them is characterized by a frequency ω_1 close to the difference $\omega_a - \omega_b$, and the other by a frequency ω_2 , close to the difference $\omega_a - \omega_c$. The perturbing Hamiltonian is

$$H_1 = eE_1x \cos \omega_1 t + eE_2x \cos \omega_2 t \quad (18B-1)$$

In matrix form, the only nonvanishing elements are taken to be $\langle a|x|b \rangle$ and $\langle a|x|c \rangle$, which can be taken to be real. We introduce the notation

$$\begin{aligned} W_{1ab} &= eE_1 \langle a|x|b \rangle \\ W_{2ab} &= eE_2 \langle a|x|b \rangle \\ W_{1ac} &= eE_1 \langle a|x|c \rangle \\ &\dots\dots\dots \end{aligned} \quad (18B-2)$$

and so on. With this notation the matrix representation of H_1 has the form

$$\begin{pmatrix} 0 & W_{1ab} \cos \omega_1 t + W_{2ab} \cos \omega_2 t & W_{1ac} \cos \omega_1 t + W_{2ac} \cos \omega_2 t \\ W_{1ab} \cos \omega_1 t + W_{2ab} \cos \omega_2 t & 0 & 0 \\ W_{1ac} \cos \omega_1 t + W_{2ac} \cos \omega_2 t & 0 & 0 \end{pmatrix} \quad (18B-3)$$

We next need to calculate $e^{iH_0 t/\hbar} H_1 e^{-iH_0 t/\hbar}$. To get this we pre-multiply the matrix for H_1

$$e^{iH_0 t/\hbar} = \begin{pmatrix} e^{i\omega_a t} & 0 & 0 \\ 0 & e^{i\omega_b t} & 0 \\ 0 & 0 & e^{i\omega_c t} \end{pmatrix} \quad (18B-4)$$

and post-multiply it by the hermitian conjugate matrix. Some algebra yields the matrix

$$V(t) = \begin{pmatrix} 0 & X & Y \\ X^* & 0 & 0 \\ Y^* & 0 & 0 \end{pmatrix} \quad (18B-5)$$

with

$$\begin{aligned} X &= (W_{1ab} \cos \omega_1 t + W_{2ab} \cos \omega_2 t) e^{i(\omega_a - \omega_b)t} \\ Y &= (W_{1ac} \cos \omega_1 t + W_{2ac} \cos \omega_2 t) e^{i(\omega_a - \omega_c)t} \end{aligned} \quad (18B-6)$$

We now again apply the rotating wave approximation, with $\omega_a - \omega_b - \omega_1 = \delta_1$, and $\omega_a - \omega_c - \omega_2 = \delta_2$ being the only terms that we keep in this approximation. This means that in X and Y we decompose the cosines, and only keep the terms below:

$$\begin{aligned} X &= \frac{1}{2} W_{1ab} e^{i\delta_1 t} \\ Y &= \frac{1}{2} W_{2ac} e^{i\delta_2 t} \end{aligned} \quad (18B-7)$$

Let us take the state vector as represented by the column vector

$$|\psi_1(t)\rangle \rightarrow \begin{pmatrix} a(t) \\ b(t) \\ c(t) \end{pmatrix} \quad (18B-8)$$

The set of equations to be solved is

$$\begin{aligned} i\hbar \frac{da(t)}{dt} &= \frac{1}{2} W_{1ab} e^{i\delta_1 t} b(t) + \frac{1}{2} W_{2ac} e^{i\delta_2 t} c(t) \\ i\hbar \frac{db(t)}{dt} &= \frac{1}{2} W_{1ab} e^{-i\delta_1 t} a(t) \\ i\hbar \frac{dc(t)}{dt} &= \frac{1}{2} W_{2ac} e^{-i\delta_2 t} a(t) \end{aligned} \quad (18B-9)$$

Let us write

$$\begin{aligned} B(t) &= e^{i\delta_1 t} b(t) \\ C(t) &= e^{i\delta_2 t} c(t) \end{aligned} \quad (18B-10)$$

In terms of these, the equations become

$$\begin{aligned} i\hbar \frac{da(t)}{dt} &= \frac{W_{1ab}}{2} B(t) + \frac{W_{2ac}}{2} C(t) \\ i\hbar \frac{dB(t)}{dt} + \hbar \delta_1 B(t) &= \frac{W_{1ab}}{2} A(t) \\ i\hbar \frac{dC(t)}{dt} + \hbar \delta_2 C(t) &= \frac{W_{2ac}}{2} A(t) \end{aligned} \quad (18B-11)$$

Let us now assume the time dependence $e^{i\Omega t}$ in all the terms. We then get from the various equations

$$\begin{aligned} -\hbar \Omega a(0) &= \frac{W_{1ab}}{2} B(0) + \frac{W_{2ac}}{2} C(0) \\ -\hbar(\Omega - \delta_1) B(0) &= \frac{W_{1ab}}{2} a(0) \\ -\hbar(\Omega - \delta_2) C(0) &= \frac{W_{2ac}}{2} a(0) \end{aligned} \quad (18B-12)$$

This leads to a cubic equation for Ω , as might have been expected. The equation reduces to

$$\Omega(\Omega - \delta_1)(\Omega - \delta_2) = \left(\frac{W_{1ab}}{2\hbar} \right)^2 (\Omega - \delta_2) + \left(\frac{W_{2ac}}{2\hbar} \right)^2 (\Omega - \delta_1) \quad (18B-13)$$

We can greatly simplify matters by assuming *perfect tuning* so that $\delta_1 = \delta_2 = 0$. In that case the equation has simple roots: $\Omega = 0, \Omega = r, \Omega = -r$, where

$$r = \sqrt{\left(\frac{W_{1ab}}{2\hbar}\right)^2 + \left(\frac{W_{2ac}}{2\hbar}\right)^2} \quad (18B-14)$$

Thus we have

$$\begin{aligned} a(t) &= a_0 + a_+ e^{irt} + a_- e^{-irt} \\ b(t) &= b_0 + b_+ e^{irt} + b_- e^{-irt} \\ c(t) &= c_0 + c_+ e^{irt} + c_- e^{-irt} \end{aligned} \quad (18B-15)$$

The nine parameters will be determined by the equations (18B-12).

We now write the solutions in terms of the eigenstates corresponding to the different eigenvalues

$$\begin{pmatrix} a(t) \\ b(t) \\ c(t) \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix} + \begin{pmatrix} a_+ \\ b_+ \\ c_+ \end{pmatrix} e^{irt} + \begin{pmatrix} a_- \\ b_- \\ c_- \end{pmatrix} e^{-irt} \quad (18B-16)$$

To find the normalized eigenstates we proceed as follows:

$\Omega = 0$: We have $a_0 = 0$, and $W_{1ab} b_0 + W_{2ac} c_0 = 0$. The normalized solutions is therefore

$$\begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix} = \begin{pmatrix} 0 \\ W_{2ac}/2\hbar r \\ -W_{1ab}/2\hbar r \end{pmatrix} \quad (18B-17)$$

This is normalized to unity since it $W_{2ac}^2 + W_{1ab}^2 = 4\hbar^2 r^2$.

For $\Omega = r$, we must satisfy

$$\begin{aligned} -\hbar r a_+ &= \frac{W_{1ab}}{2} b_+ + \frac{W_{2ac}}{2} c_+ \\ -\hbar r b_+ &= \frac{W_{1ab}}{2} a_+ \\ -\hbar r c_+ &= \frac{W_{2ac}}{2} a_+ \end{aligned} \quad (18B-18)$$

The first equation is automatically satisfied if the other two are, which is to be expected, since a_+ is to be determined by normalization. A little algebra shows that

$$\begin{pmatrix} a_+ \\ b_+ \\ c_+ \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -W_{1ab}/2\hbar r \\ -W_{2ac}/2\hbar r \end{pmatrix} \quad (18B-19)$$

The case for $\Omega = -r$ is easily solved by just changing the sign of r . We get

$$\begin{pmatrix} a_- \\ b_- \\ c_- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ W_{1ab}/2\hbar r \\ W_{2ac}/2\hbar r \end{pmatrix} \quad (18B-20)$$

It is easy to check that the three eigenvectors are mutually orthogonal.

The general solution is

$$\begin{pmatrix} a(t) \\ b(t) \\ c(t) \end{pmatrix} = \alpha_0 \begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix} + \alpha_+ \begin{pmatrix} a_+ \\ b_+ \\ c_+ \end{pmatrix} e^{irt} + \alpha_- \begin{pmatrix} a_- \\ b_- \\ c_- \end{pmatrix} e^{-irt} \quad (18B-21)$$

and the coefficients $(\alpha_0, \alpha_+, \alpha_-)$ are determined by the initial conditions. At time $t = 0$ we have

$$\begin{aligned} a(0) &= \frac{1}{\sqrt{2}} \alpha_+ + \frac{1}{\sqrt{2}} \alpha_- \\ b(0) &= \frac{W_{2ac}}{2\hbar r} \alpha_0 - \frac{W_{1ab}}{2\sqrt{2}\hbar r} \alpha_+ + \frac{W_{1ab}}{2\sqrt{2}\hbar r} \alpha_- \\ c(0) &= \frac{W_{1ab}}{2\hbar r} \alpha_0 - \frac{W_{2ac}}{2\sqrt{2}\hbar r} \alpha_+ + \frac{W_{2ac}}{2\sqrt{2}\hbar r} \alpha_- \end{aligned} \quad (18B-22)$$

Dark States

Consider a configuration in which the $|b\rangle$ and the $|c\rangle$ states lie close together (Fig. 18-3a) and

$$W_{1ab} = W_{2ac} \quad (18B-23)$$

We now take our initial state to be

$$|\psi_1(0)\rangle = \frac{1}{\sqrt{2}} (|b\rangle - |c\rangle) \quad (18B-24)$$

This implies that $a(0) = 0$; $b(0) = -c(0) = 1/\sqrt{2}$. We now solve for $(\alpha_0, \alpha_+, \alpha_-)$, and get

$$\begin{aligned} \alpha_0 &= \frac{W_{1ab} + W_{2ac}}{2\sqrt{2}\hbar r} = \frac{W_{1ab}}{\sqrt{2}\hbar r} \\ \alpha_+ &= -\alpha_- = \frac{W_{2ac} - W_{1ab}}{4\hbar r} = 0 \end{aligned} \quad (18B-25)$$

This implies that $a(t) = 0$ for all times. Thus the state $|a\rangle$ is *never* excited, and is therefore called a *dark state*. The reason it is inaccessible is that the amplitudes for exciting from the $|b\rangle$ and $|c\rangle$ states interfere destructively.

Electromagnetically Induced Transparency

Consider, next, a situation in which the $|a\rangle$ and $|b\rangle$ states are strongly coupled by an electromagnetic field, while $|a\rangle$ and $|c\rangle$ are weakly coupled. What this implies is that

$$|W_{1ab}| \gg |W_{2ac}| \quad (18B-26)$$

One can show that under these circumstances the state $|a\rangle$ is very unlikely to be excited, and this means that photons cannot be absorbed by a $|c\rangle \rightarrow |a\rangle$ transition.

We take for our initial condition $a(0) = b(0) = 0$. This implies that

$$\begin{aligned} \alpha_- &= -\alpha_+ \\ \alpha_+ &= \frac{W_{2ac}}{\sqrt{2} W_{1ab}} \alpha_0 \end{aligned} \quad (18B-27)$$

Furthermore, since $a(0) = b(0) = 0$, we must necessarily have $|c(0)| = 1$. Since

$$c(0) = \frac{2\hbar r}{W_{1ab}} \alpha_0$$

W-90 Supplement 18-B Details of the Three-Level System

we may choose for convenience

$$\alpha_0 = \frac{W_{1ab}}{2\hbar r} \quad (18B-28)$$

Now, at a later time t , we get

$$a(t) = \frac{\alpha_+}{\sqrt{2}} (e^{irt} - e^{-irt}) = i \frac{W_{2ac}}{2\hbar r} \sin rt \quad (18B-29)$$

Thus the probability of finding the system in the state $|a\rangle$ at a later time is

$$P_a(t) = |a(t)|^2 = \frac{W_{2ac}^2}{W_{2ac}^2 + W_{1ab}^2} \sin^2 rt \quad (18B-30)$$

Because of the condition (18B-26) the probability of exciting the state $|a\rangle$ is very small. This, however, implies that photons cannot be absorbed through the mechanism of exciting state $|a\rangle$, so the material becomes *transparent* at the frequency corresponding to this energy difference.

Chapter 19

Collision Theory

Atomic and molecular structure was largely explored through spectroscopy. When it comes to trying to understand nuclear forces and the laws that govern the interactions of elementary particles, the only technique available is that of scattering a variety of particles by a variety of targets. In some sense, spectroscopy is also a form of “scattering.” The atom in the ground state is excited by some projectile (it may be electrons in a discharge tube or collisions with other target particles, as in heating up of the gas), and then an outgoing photon is observed, with the atom going into the ground state again, or possibly another excited state. We do not usually describe these processes as “collision processes” because the atom has very well-defined energy levels, in which it stays for times that are enormously long compared to collision times,¹ so that it is possible to separate the “decay” from the excitation process. In particular, the characteristics of the decay are not sensitive to the particular mode of excitation. In nuclei and also in elementary particles, there exist levels, but frequently the lifetime is not sufficiently long to warrant a separation into excitation and decay, especially since accompanying the “resonant” scattering there is also nonresonant “background” scattering, and the disentangling of the two is sometimes complicated. In this chapter we will therefore discuss the process as a whole.

19-1 COLLISION CROSS SECTION

The ideal way to talk about scattering is to formulate equations that describe exactly what happens: An incident particle, described by a wave packet, approaches the target. The wave packet must be spatially large, so that it does not spread appreciably during the experiment, and it must be large compared with the target particle, but small compared with the dimensions of the laboratory; that is, it must not simultaneously overlap the target and detector. The lateral dimensions are, in fact, determined by the beam size in the accelerator. There follows an interaction with the target, and finally we see two wave packets: One continues in the forward direction, describing the unscattered part of the beam, and the other flies off at some angle and describes the scattered particles. The number of particles scattered into a given solid angle per unit time and unit incident flux is defined to be the *differential scattering cross section*. We will not follow this approach directly, but will instead use some of the material developed in Supplement 8-B [www.wiley.com/college/gasiorowicz] to obtain the differential cross section. We will, however, keep the wave-packet treatment in mind as we interpret our formal results.

¹Recall that the lifetime of a $2p$ hydrogen state is 1.6×10^{-9} s, which is large compared to the characteristic time $a_0/c \approx 2 \times 10^{-17}$ s.

In our discussion of the continuum solutions of the Schrödinger equation in Supplement 8-B [www.wiley.com/college/gasiorowicz] we concluded that: (a) A solution of the Schrödinger equation in the absence of a potential is the plane wave form $e^{i\mathbf{k}\cdot\mathbf{r}}$, which describes a flux

$$\mathbf{j} = \frac{\hbar}{2im} (\psi^* \nabla \psi - \psi \nabla \psi^*) = \frac{\hbar \mathbf{k}}{m} \quad (19-1)$$

If we choose $\mathbf{k}$ to define the z -axis, then the large r behavior of this solution may be written in the form of an incoming + an outgoing spherical wave

$$e^{i\mathbf{k}\cdot\mathbf{r}} \Rightarrow \frac{i}{2k} \sum_{l=0}^{\infty} (2l+1) i^l \left[\frac{e^{-i(kr-l\pi/2)}}{r} - \frac{e^{i(kr-l\pi/2)}}{r} \right] P_l(\cos \theta) \quad (19-2)$$

(b) The conservation of particles forces us to the conclusion that the presence of a radial potential can only alter this to a function, whose asymptotic form is

$$\psi(\mathbf{r}) \Rightarrow \frac{i}{2k} \sum_{l=0}^{\infty} (2l+1) i^l \left[\frac{e^{-i(kr-l\pi/2)}}{r} - S_l(k) \frac{e^{i(kr-l\pi/2)}}{r} \right] P_l(\cos \theta) \quad (19-3)$$

subject to²

$$|S_l(k)| = 1 \quad (19-4)$$

The asymptotic form (19-3) may be rewritten, with the help of (19-2), as

$$\psi(\mathbf{r}) \Rightarrow e^{i\mathbf{k}\cdot\mathbf{r}} + \left[\sum_{l=0}^{\infty} (2l+1) \frac{S_l(k) - 1}{2ik} P_l(\cos \theta) \right] \frac{e^{ikr}}{r} \quad (19-5)$$

corresponding to a plane wave + an outgoing spherical wave. Note that we are working with the effective one-particle Schrödinger equation, so that m is the reduced mass and θ is the center of mass angle between the direction of $\mathbf{k}$ (the z -axis) and the asymptotic point $\mathbf{r}$, where presumably the counter will be set up. When the target is much more massive than the projectile, there is no distinction between the laboratory angle and the center-of-mass angle. Note also that we could, of course, have set up a solution that has the asymptotic form of a plane wave + an incoming spherical wave, since it is the first term in (19-3) that could be modified by a coefficient satisfying (19-4). However, the solution that describes the scattering is the one involving the outgoing wave. Let us calculate the flux for the asymptotic solution (19-5).

$$\mathbf{j} = \frac{\hbar}{2im} \left\{ \left[e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta) \frac{e^{ikr}}{r} \right]^* \nabla \left[e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta) \frac{e^{ikr}}{r} \right] - \text{complex conjugate} \right\} \quad (19-6)$$

where we have defined

$$f(\theta) = \sum_{l=0}^{\infty} (2l+1) f_l(k) P_l(\cos \theta) \quad (19-7)$$

with

$$f_l(k) = [S_l(k) - 1]/2ik \quad (19-8)$$

²See the discussion following (8B-3). $S_l(k)$ is the standard notation for $e^{2\delta_l(k)}$ defined in (8B-6). See Supplement 8-B [www.wiley.com/college/gasiorowicz].

Calculating the gradient gives

$$\begin{aligned} \mathbf{j} &= \frac{\hbar}{2im} \left\{ \left[e^{-i\mathbf{k}\cdot\mathbf{r}} + f^*(\theta) \frac{e^{-ikr}}{r} \right] \left[i\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}} + \hat{i}_\theta \frac{1}{r} \frac{\partial f(\theta)}{\partial \theta} \frac{e^{ikr}}{r} \right. \right. \\ &\quad \left. \left. + \hat{i}_r f(\theta) \left(ik \frac{e^{ikr}}{r} - \frac{e^{ikr}}{r^2} \right) \right] - \text{complex conjugate} \right\} \\ &= \frac{\hbar}{2im} \left\{ i\mathbf{k} + i\mathbf{k}f^*(\theta) \frac{e^{-ikr(1-\cos\theta)}}{r} + ik\hat{i}_r f(\theta) \frac{e^{ikr(1-\cos\theta)}}{r} + ik\hat{i}_r |f(\theta)|^2 \frac{1}{r^2} \right. \\ &\quad \left. - \hat{i}_r f(\theta) \frac{e^{ikr(1-\cos\theta)}}{r^2} + \hat{i}_\theta \frac{\partial f(\theta)}{\partial \theta} \frac{e^{ikr(1-\cos\theta)}}{r^2} - \text{complex conjugate} \right\} \end{aligned}$$

We have used $\mathbf{k} \cdot \mathbf{r} = kr \cos \theta$, in the exponential factors. $\hat{i}_r$ is the unit vector in the $\mathbf{r}$ direction.

In carrying out this calculation we have left out the $1/r^3$ terms, since these are dominated by the $1/r^2$ terms for large r . We will be interested in the flux at a detector, which is located at a distance r from the origin, where the potential that gives rise to the scattering is localized, and thus in large r . Thus the flux is

$$\begin{aligned} \mathbf{j} &= \frac{\hbar\mathbf{k}}{m} + \frac{\hbar k}{m} \hat{i}_r |f(\theta)|^2 \frac{1}{r^2} \\ &\quad + \frac{\hbar\mathbf{k}}{2m} \frac{1}{r} \left[f^*(\theta) e^{-ikr(1-\cos\theta)} + f(\theta) e^{ikr(1-\cos\theta)} \right] \\ &\quad + \frac{\hbar k}{2m} \frac{\hat{i}_r}{r} \left[f^*(\theta) e^{-ikr(1-\cos\theta)} + f(\theta) e^{ikr(1-\cos\theta)} \right] \\ &\quad - \frac{\hbar}{2im} \frac{\hat{i}_r}{r^2} \left[f(\theta) e^{ikr(1-\cos\theta)} - f^*(\theta) e^{-ikr(1-\cos\theta)} \right] \\ &\quad + \frac{\hbar}{2im} \frac{\hat{i}_\theta}{r^2} \left[\frac{\partial f(\theta)}{\partial \theta} e^{ikr(1-\cos\theta)} - \frac{\partial f^*(\theta)}{\partial \theta} e^{-ikr(1-\cos\theta)} \right] \end{aligned} \quad (19-9)$$

This rather involved expression simplifies considerably when we consider that $\theta \neq 0$, since one never does a scattering experiment directly in the forward direction,³ and that in a measurement one always integrates the flux over a small but finite solid angle. Thus in the last four terms of this expression we should replace $e^{ikr(1-\cos\theta)}$ by

$$\int \sin \theta \, d\theta \, d\phi \, g(\theta, \phi) e^{ikr(1-\cos\theta)} \quad (19-10)$$

where $g(\theta, \phi)$ is some sort of smooth, localized acceptance function for the counter. Now, as $r \rightarrow \infty$, we have an integral over a product of a smooth function and an extremely rapidly varying one, and this vanishes faster than any power of $1/r$. This is what is known in the mathematical literature as the Riemann–Lebesgue lemma, and it is illustrated in Problem 7. Thus, only the first two terms remain, so that

$$\mathbf{j} = \frac{\hbar\mathbf{k}}{m} + \frac{\hbar k}{m} \hat{i}_r \frac{|f(\theta)|^2}{r^2} \quad (19-11)$$

³How could one tell scattered from unscattered particles?

In the absence of a potential, only the first term is there; it represents the incident flux. In a wave-packet treatment, $\hbar \mathbf{k}/m$ would be multiplied by a function that defines the lateral dimensions of the beam. Thus, if we ask for the *radial flux*, $\hat{\mathbf{i}}_r \cdot \mathbf{j}$, then that term gives a contribution $\hbar \mathbf{k} \cdot \hat{\mathbf{i}}_r/m = \hbar k \cos \theta/m$, but only within a finite region of the z -axis (see Fig. 19-1). Since the counter is put outside of that region, this first term does not contribute to the radial flux in the asymptotic region, so that only the second term in (19-11) contributes, and

$$\mathbf{j} \cdot \hat{\mathbf{i}}_r = \frac{\hbar k}{m} \cdot \frac{|f(\theta)|^2}{r^2} \quad (19-12)$$

Thus the number of particles crossing the area that subtends a solid angle $d\Omega$ at the origin (the target) is

$$\mathbf{j} \cdot \hat{\mathbf{i}}_r dA = \frac{\hbar k}{m} \cdot \frac{|f(\theta)|^2}{r^2} r^2 d\Omega \quad (19-13)$$

Notice that the factor of r^2 drops out. This justifies the dropping of the $1/r^3$ terms in (19-9), since these would contribute terms of order $1/kr$ to the number of particles. The differential cross section is this number, divided by the incident flux, $\hbar k/m$; that is,

$$d\sigma = |f(\theta)|^2 d\Omega \quad (19-14)$$

If the potential has spin dependence, there may be an azimuthal dependence, so that more generally,

$$\frac{d\sigma}{d\Omega} = |f(\theta, \phi)|^2 \quad (19-15)$$

The total cross section is given by

$$\sigma_{\text{tot}}(k) = \int d\Omega \frac{d\sigma}{d\Omega} \quad (19-16)$$

If we now use $f(\theta)$ as expressed in terms of $S_l(k)$, and express the latter in terms of the phase shift $S_l(k) = e^{2i\delta_l(k)}$, so that

$$f(\theta) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l(k)} \sin \delta_l(k) P_l(\cos \theta) \quad (19-17)$$

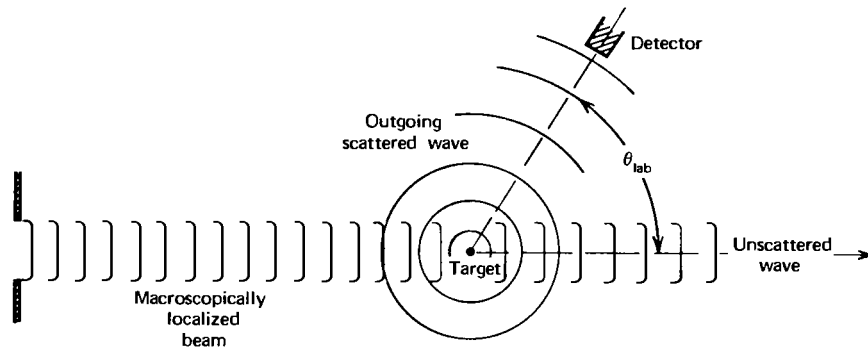


Figure 19-1 Schematic layout for scattering experiment. The scattering angle is the laboratory angle.

then

$$\sigma_{\text{tot}} = \int d\Omega \left[\frac{1}{k} \sum_l (2l+1) e^{i\delta_l(k)} \sin \delta_l(k) P_l(\cos \theta) \right] \left[\frac{1}{k} \sum_{l'} (2l'+1) e^{-i\delta_{l'}(k)} \sin \delta_{l'}(k) P_{l'}(\cos \theta) \right]$$

and using

$$\int d\Omega P_l(\cos \theta) P_{l'}(\cos \theta) = \frac{4\pi}{2l+1} \delta_{ll'} \quad (19-18)$$

we get

$$\sigma_{\text{tot}} = \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l(k) \quad (19-19)$$

It is an interesting fact that

$$\begin{aligned} \text{Im } f(0) &= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) \text{Im}[e^{i\delta_l(k)} \sin \delta_l(k)] P_l(1) \\ &= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l(k) = \frac{k}{4\pi} \sigma_{\text{tot}} \end{aligned} \quad (19-20)$$

This relation is known as the *optical theorem* and it is true even when inelastic processes can occur, as they do in nuclear and particle physics scattering processes. It is a very useful relation and in wave language it follows from the fact that the total cross section represents the removal of flux from the incident beam. Such a removal can only occur as a result of destructive interference, and the latter can only occur between the incident wave and the elastically scattered wave in the forward direction. This explains why $f(0)$ appears linearly.

This hand-waving argument does not explain why it is the imaginary part that is involved, but this can be shown to be generally true.⁴

Elastic and Inelastic Scattering

The requirement that $|S_l(k)| = 1$ followed from conservation of flux. Actually, in many scattering experiments there is *absorption* of the incident beam; the target may merely get excited, or change its state, or another particle may emerge. Under these circumstances our discussion is unchanged except that

$$S_l(k) = \eta_l(k) e^{2i\delta_l(k)} \quad (19-21)$$

is to be used, with

$$0 \leq \eta_l(k) \leq 1 \quad (19-22)$$

because we are dealing with absorption. The partial wave scattering amplitude is now

$$f_l(k) = \frac{S_l(k) - 1}{2ik} = \frac{\eta_l(k) e^{2i\delta_l(k)} - 1}{2ik} = \frac{\eta_l \sin 2\delta_l}{2k} + i \frac{1 - \eta_l \cos 2\delta_l}{2k} \quad (19-23)$$

⁴See L. I. Schiff, *Progr. Theor. Phys.* (Kyoto), **11**, 288 (1954).

and the total *elastic* cross section is

$$\begin{aligned}\sigma_{el} &= 4\pi \sum_l (2l+1) |f_l(k)|^2 \\ &= 4\pi \sum_l (2l+1) \frac{1 - \eta_l^2 - 2\eta_l \cos 2\delta_l}{4k^2}\end{aligned}\quad (19-24)$$

There is also a cross section for the *inelastic* processes. Since we do not specify what the inelastic processes consist of, we can only talk about the *total inelastic cross section*, which describes the loss of flux. If we look at a particular term in (19-3), the inward radial flux carried by

$$\frac{i}{2k} \frac{e^{-ikr}}{r} P_l(\cos \theta)$$

is

$$\left(\frac{\hbar k}{m}\right) \left[\frac{4\pi}{(2k)^2} \right]$$

(Recall that $Y_0 = P_0(\cos \theta)/\sqrt{4\pi}$). The outward radial flux is $(\hbar k/m)(|S_l(k)|^2 4\pi/4k^2)$, so that the net flux lost is $(\hbar k/m)(\pi/k^2)[1 - \eta_l^2(k)]$ for each l -value. Hence, dividing by the incident flux, we get

$$\sigma_{inel} = \frac{\pi}{k^2} \sum_l (2l+1) [1 - \eta_l^2(k)] \quad (19-25)$$

Thus the total cross section is

$$\begin{aligned}\sigma_{tot} &= \sigma_{el} + \sigma_{inel} \\ &= \frac{\pi}{k^2} \sum_l (2l+1) (1 + \eta_l^2 - 2\eta_l \cos 2\delta_l + 1 - \eta_l^2) \\ &= \frac{2\pi}{k^2} \sum_l (2l+1) (1 - \eta_l \cos 2\delta_l)\end{aligned}\quad (19-26)$$

It also follows from (19-23) that

$$\begin{aligned}\text{Im } f(0) &= \sum_l (2l+1) \text{Im } f_l(k) \\ &= \sum_l (2l+1) \frac{1 - \eta_l \cos 2\delta_l}{2k} = \frac{k}{4\pi} \sigma_{tot}\end{aligned}\quad (19-27)$$

so that the optical theorem is indeed satisfied.

If $\eta_l(k) = 1$, we have no absorption, and the inelastic cross section vanishes. When $\eta_l(k) = 0$ we have total absorption. Nevertheless there is still elastic scattering in that partial wave. This becomes evident in *scattering by a black disc*. The black disc is described as follows: (a) it has a well-defined edge; and (b) it is totally absorbing. Since we will consider scattering for short wavelength—that is, large k -values—condition (a) specifies that we only consider partial waves $l \leq L$, where

$$L = ka \quad (19-28)$$

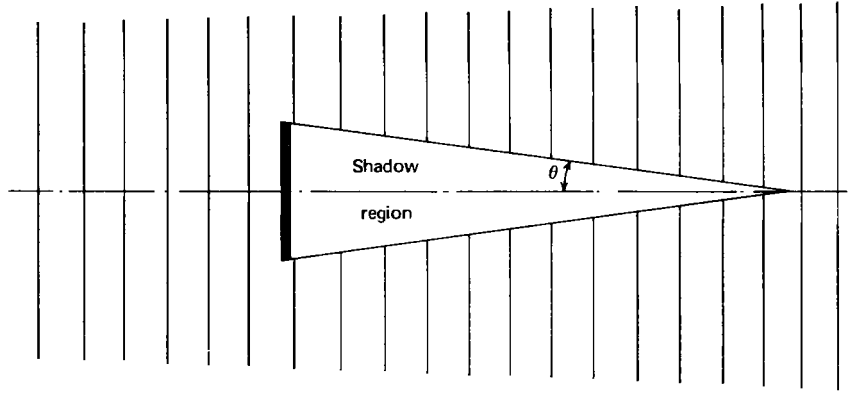


Figure 19-2 Black disc scattering and the shadow effect.

and a is the radius of the disc. Condition (b) specifies that $\eta_l(k) = 0$ for the relevant values of $l \leq L$. Thus

$$\sigma_{\text{inel}} = \frac{\pi}{k^2} \sum_{l=0}^L (2l+1) = \frac{\pi}{k^2} L^2 = \pi a^2 \quad (19-29)$$

and

$$\sigma_{\text{el}} = \frac{\pi}{k^2} \sum_{l=0}^L (2l+1) = \pi a^2 \quad (19-30)$$

so that the total cross section is

$$\sigma_{\text{tot}} = \sigma_{\text{el}} + \sigma_{\text{inel}} = 2\pi a^2 \quad (19-31)$$

The result looks peculiar; on purely classical grounds we might perhaps expect that the total cross section cannot exceed the area presented by the disc; we might also expect to see no elastic scattering when there is total absorption. This is wrong; the absorptive disc takes flux proportion to πa^2 out of the incident beam (Fig. 19-2), and this leads to a shadow behind the disc. Far away, however, the shadow gets filled in—far enough away you cannot “see” the disc—and the only way in which this can happen is through the diffraction of some of the incident wave at the edge of the disc. The amount of incident wave that must be diffracted is the same amount as was taken out of the beam to make the shadow. Thus the elastically scattered flux must also be proportional to πa^2 . The elastic scattering that accompanies absorption is called *shadow scattering* for the above reason. It is strongly peaked forward. The angle to which it is confined can be estimated from the uncertainty principle: An uncertainty in the lateral direction of magnitude a will be accompanied by an uncontrolled lateral momentum transfer of magnitude $p_{\perp} \sim \hbar/a$. This, however, is equal to $p\theta$, so that

$$\theta \sim \frac{\hbar}{ap} \sim \frac{1}{ak} \quad (19-32)$$

This agrees with the optical result $\theta \sim \lambda/a$. These features are observed both in nuclear scattering and in particle scattering at high energies, since the central region of nuclei and of protons is strongly absorptive, and the edges of these objects are moderately sharp. (See Fig. 19-3.)

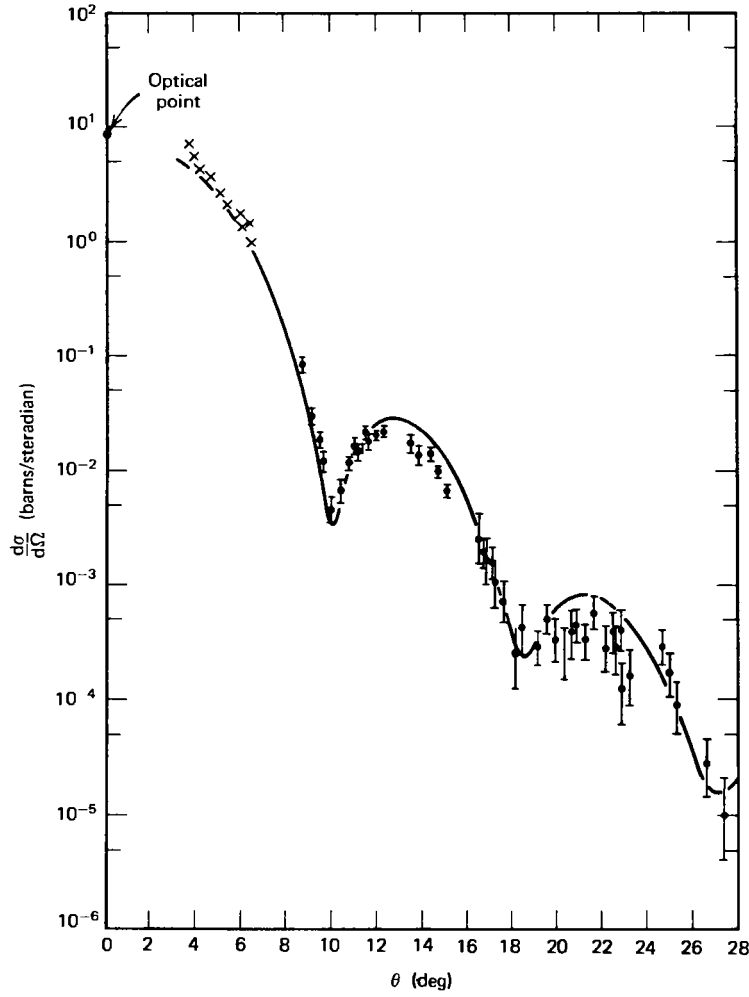


Figure 19-3 Angular distribution of 1000 MeV (1 GeV) protons scattered by ^{16}O nuclei. The angular distribution shows the dips that characterize diffraction scattering. The departures from the shape of Fraunhofer scattering in optics is due to the fact that nuclei are not sharp, nor are they totally absorbing. The curve is the result of a theoretical calculation that takes these effects into account. (From H. Palevsky et al., *Phys. Rev. Lett.* **18**, 1200 (1967), by permission.)

19-2 SCATTERING AT LOW ENERGIES

The phase shift expansion (19-17) may be used to express the differential cross section in terms of the phase shifts

$$\frac{d\sigma}{d\Omega} = \frac{1}{k^2} \left| \sum_l (2l+1) e^{i\delta_l(k)} \sin \delta_l(k) P_l(\cos \theta) \right|^2 \quad (19-33)$$

We expect, on grounds of correspondence with classical theory, that the angular momentum involved in the scattering is bounded by pa where p is the center-of-mass momentum and a is the range of the forces. Thus we expect that

$$l \lesssim \frac{pa}{\hbar} = ka \quad (19-34)$$

With the sum in (19-33) limited, one can try, by fitting the differential cross section measured at a number of angles to form like

$$\frac{d\sigma}{d\Omega} = \sum_{n=0}^N A_n (\cos \theta)^n \quad (19-35)$$

to determine the phase shifts for a finite number of l -values. There are ambiguities—for example, the cross section is unaltered when all the phase shifts change their sign—but these can be resolved with the help of theory, continuity from low energies, and other tricks of the trade. The hope is that one can learn something about the interaction from the phase shifts, which form empirical data somewhat closer to the theory than the cross sections do.

The connection between the phase shifts $\delta_l(k)$ and the potential $V(r)$ is via the Schrödinger equation; the radial equation will have a solution that asymptotically behaves as

$$R_l(r) \sim \frac{1}{r} \sin \left[kr - \frac{l\pi}{2} + \delta_l(k) \right] \quad (19-36)$$

aside from an amplitude factor in front. Thus, given $V(r)$, a straightforward way to calculate $\delta_l(k)$ is to integrate the radial equation numerically to values of r that are far out of the range of the potential, and to examine the asymptotic behavior. This is, in fact, what one does, but this does not give us any insight into the properties of the phase shifts. To learn more about the phase shifts, we consider the square well potential. We found in Supplement 8-B [www.wiley.com/college/gasiorowicz], Eq. (8B-13), that

$$\tan \delta_l(k) = -\frac{C}{B} \quad (19-37)$$

where the ratio is obtained by matching the internal to the external radial wave function (8B-12)

$$\kappa \frac{j_l'(\kappa a)}{j_l(\kappa a)} = k \frac{j_l'(ka) + (C/B) n_l'(ka)}{j_l(ka) + (C/B) n_l(ka)} \quad (19-38)$$

In this equation

$$\kappa^2 = \frac{2m}{\hbar^2} (E + V_0) \quad k^2 = \frac{2mE}{\hbar^2} \quad (19-39)$$

and the ' denotes differentiation, with respect to the argument. $V_0 > 0$ for an attractive potential. Thus

$$\tan \delta_l(k) = \frac{k j_l'(ka) j_l(\kappa a) - \kappa j_l(ka) j_l'(\kappa a)}{\kappa n_l'(ka) j_l(\kappa a) - \kappa n_l(ka) j_l'(\kappa a)} \quad (19-40)$$

This is not a particularly transparent expression, but it simplifies in some limiting cases.

(a) Consider the case that

$$ka \ll l \quad (19-41)$$

We do not insist that $\kappa a \ll l$. With the help of the formulas (8-65) and (8-66) we get

$$\tan \delta_l(k) \approx \frac{2l+1}{[1.3.5 \dots (2l+1)]^2} (ka)^{2l+1} \frac{l j_l(\kappa a) - \kappa a j_l'(\kappa a)}{(l+1) j_l(ka) + \kappa a j_l'(ka)} \quad (19-42)$$

after a little algebra. One can show that for large l , this drops faster than e^{-l} even if $ka \gg 1$. The behavior

$$\tan \delta_l(k) \sim k^{2l+1} \quad (19-43)$$

for $ka \rightarrow 0$ is not restricted to the square well potential, but is true for all reasonably smooth potentials. It is a consequence of the centrifugal barrier, which keeps waves of energy far below the barrier from feeling the effect of the potential.

(b) For certain values of the energy, the denominator in (19-40) will vanish, so that at these energies the phase shift passes through $\pi/2$, or more generally through $(n + 1/2)\pi$. When the phase shift is $\pi/2$, then the partial wave cross section

$$\sigma_l(k) = \frac{4\pi(2l+1)}{k^2} \sin^2 \delta_l(k) \quad (19-44)$$

has the largest possible value. One says that when $\tan \delta_l(k)$ rises rapidly to infinity and continues rising from $-\infty$, we have *resonant scattering*. To justify this terminology, and explain when resonant scattering occurs, let us consider a very deep potential, and also l large, so that

$$\kappa a \gg l \gg ka \quad (19-45)$$

We may then use (19-42) for $\tan \delta_l(k)$, and this will become infinite when

$$(l+1)j_l(\kappa a) + \kappa a j_l'(\kappa a) = 0 \quad (19-46)$$

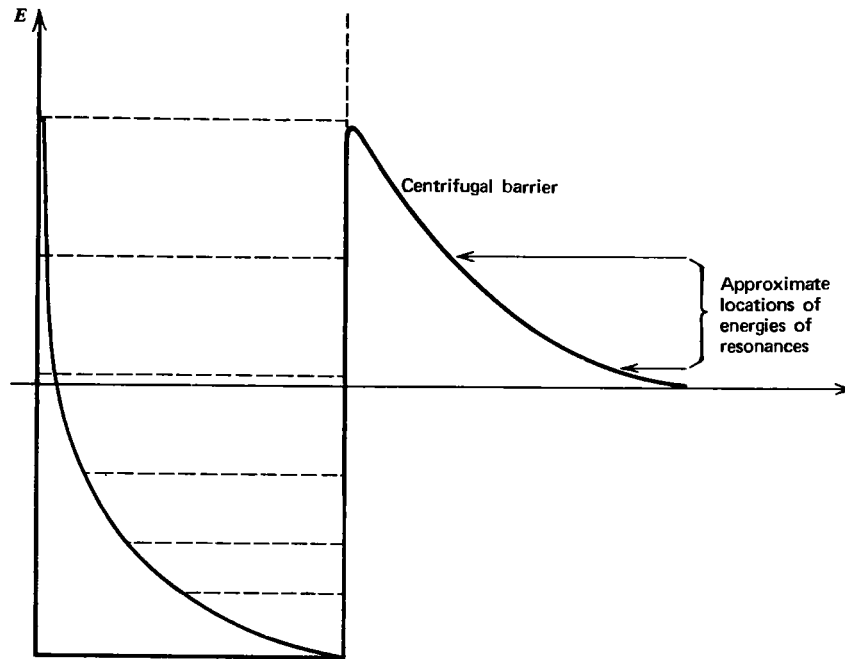


Figure 19-4 Sketch showing the square well potential with the centrifugal barrier tail. The dashed lines represent the energy levels in an infinite square well of range a , and the approximate locations of the scattering resonance energies are indicated on the right. The lower one will be much sharper than the upper one.

Since $\kappa a \gg l$, this condition is approximately equivalent to

$$\frac{(l+1)}{\kappa a} \cos\left(\kappa a - \frac{1}{2}\pi\right) - \sin\left(\kappa a - \frac{1}{2}\pi\right) = 0$$

that is,

$$\tan\left(\kappa a - \frac{1}{2}\pi\right) \approx \frac{l+1}{\kappa a} \quad (19-47)$$

Since the right side is very small, the resonance condition is

$$\kappa a - \frac{1}{2}\pi \approx n\pi + \frac{l+1}{\kappa a} \quad (19-48)$$

Now this is just the condition (8-73') for the existence of discrete levels in a three-dimensional box, so that resonant scattering occurs when the incident energy is just such as to match an energy level. Since $E > 0$, these levels are not really bound states. As Fig. 19-4 indicates, these are levels that would be bound states if the barrier were infinitely thick. It is not, but a particle being scattered at just the right energy still “knows” that there is a virtual level there.

In Supplement 15-A [www.wiley.com/college/gasiorowicz] we discuss the scattering of a photon at an energy corresponding to a state that would be stationary in the absence of coupling to the radiation field. There we have the same situation, and we also see a resonant behavior.

The Breit–Wigner Formula

As (19-42) shows, the phase shift is very tiny for ka small. Nevertheless, as ka changes and goes through the resonance, δ_l rises very rapidly, increasing by π ; thus the partial wave cross section (19-44) will exhibit a very sharp peak at the resonant energy. This behavior (Fig. 19-5) is very similar to the cross section for the scattering of electrons by He^+ at the energy corresponding to the $(2s)^2$ excited state (Fig. 14-4). In the neighborhood of the resonant energy, the phase shift rises through $\pi/2$ very rapidly. We may represent this behavior by

$$\tan \delta_l \approx \frac{\gamma(ka)^{2l+1}}{E - E_{\text{res}}} \quad (19-49)$$

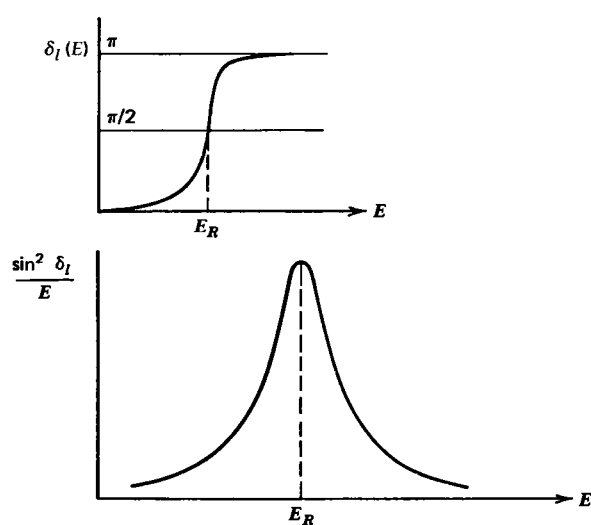


Figure 19-5 The partial wave cross section corresponding to the phase shift sketched in the upper insert.

This leads to the partial wave cross section

$$\sigma_l = \frac{4\pi(2l+1)}{k^2} \frac{\tan^2 \delta_l}{1 + \tan^2 \delta_l} = \frac{4\pi(2l+1)}{k^2} \frac{[\gamma(ka)^{2l+1}]^2}{(E - E_{\text{res}})^2 + [\gamma(ka)^{2l+1}]^2} \quad (19-50)$$

This is the well-known *Breit-Wigner formula* for resonant cross sections. Again, the behavior is not a peculiarity of the square well potential, but is characteristic of all potentials that have a shape such that metastable states can simulate bound states above $E = 0$ in it. We just note for completeness that

$$\begin{aligned} f_l(k) &= \frac{e^{2i\delta_l(k)} - 1}{2ik} = \frac{\frac{1 + i \tan \delta_l}{1 - i \tan \delta_l} - 1}{2ik} \\ &= \frac{\tan \delta_l}{k(1 - i \tan \delta_l)} = \frac{\gamma(ka)^{2l+1}/k}{E - E_{\text{res}} - i\gamma(ka)^{2l+1}} \end{aligned} \quad (19-51)$$

If there is nonresonant scattering that is appreciable, then the scattering amplitude is of the form

$$f_l(k) = f_l^{\text{res}}(k) + f_l^{\text{nonres}}(k) \quad (19-52)$$

S-Wave Scattering for a Square Well

At low energies, the scattering is primarily in S -states, so that we may concentrate on $l = 0$. It is simpler to derive the phase shift directly than to work out (19-40). The solution inside the well that is regular at $r = 0$ is

$$u(r) = rR(r) = C \sin \kappa r \quad (19-53)$$

and this is to be matched onto

$$u(r) = \sin(kr + \delta) \quad (19-54)$$

the solution outside the well. The continuity of $(1/u)(du/dr)$ at $r = a$ implies that

$$\kappa \cot \kappa a = k \cot(ka + \delta)$$

that is,

$$\tan \delta = \frac{(k/\kappa) \tan \kappa a - \tan ka}{1 + (k/\kappa) \tan \kappa a \tan ka} \quad (19-55)$$

Note that if we define

$$\tan qa = \frac{k}{\kappa} \tan \kappa a$$

then

$$\tan \delta = \frac{\tan qa - \tan ka}{1 + \tan qa \tan ka} = \tan(qa - ka)$$

that is,

$$\delta = \tan^{-1} \left(\frac{k}{\kappa} \tan \kappa a \right) - ka \quad (19-56)$$

We have, following (19-39),

$$(\kappa a)^2 = (ka)^2 + \frac{2mV_0 a^2}{\hbar^2} \quad (19-57)$$

with $V_0 > 0$ for an attractive potential. Thus, at very low energies, using $\tan x \approx x$ for $x \ll 1$, we get

$$\tan \delta \approx \delta \approx ka \left(\frac{\tan \kappa a}{\kappa a} - 1 \right) \quad (19-58)$$

When κa goes through $\pi/2$ (we imagine that we are slowly deepening the potential well), which is just the condition that the well be deep enough for a bound state to develop [cf. eq. (4-52)], then $\tan \kappa a \rightarrow \infty$ and (19-55) shows that

$$\tan \delta = \frac{1}{\tan \kappa a} \rightarrow \infty \quad (19-59)$$

that is, δ goes through $\pi/2$. In a sense, a bound state at zero energy is like a resonance.

As the well becomes a little deeper, we again have $\tan \delta \sim O(ka)$, and continuity demands that the branch is such that

$$\begin{aligned} \delta &\approx ka \left(\frac{\tan \kappa a}{\kappa a} - 1 \right) && \text{(no bound state)} \\ \delta &\approx \pi + ka \left(\frac{\tan \kappa a}{\kappa a} - 1 \right) && \text{(with bound state)} \end{aligned} \quad (19-60)$$

As the potential becomes still deeper, a second bound state can appear, κa goes through $3\pi/2$, and we have $\delta \approx 2\pi + ka[(\tan \kappa a/\kappa a) - 1]$, and so on. There is a general result known as Levinson's theorem, which states

$$\delta(0) - \delta(\infty) = N_B \pi \quad (19-61)$$

where N_B is the number of bound states, and (19-60) is an example of it.

Connection between Scattering Amplitude and Binding Energy

At very low energies the cross section only has the $l = 0$ contribution to it, and it is

$$\sigma \cong \frac{4\pi}{k^2} (ka)^2 \left(\frac{\tan \kappa a}{\kappa a} - 1 \right)^2 = 4\pi a^2 \left(\frac{\tan \kappa a}{\kappa a} - 1 \right)^2 \quad (19-62)$$

that is, it is a constant. There will, of course, be a correction of order $(ka)^2$ to this result. If we consider neutron-proton scattering, then we know that the potential must be such as to give the right binding energy of the deuteron. If we let

$$E = -\frac{\hbar^2 \alpha^2}{2m}$$

and

$$\kappa = \sqrt{-\alpha^2 + \frac{2mV_0}{\hbar^2}}$$

(effectively $k^2 = -\alpha^2$ for the bound-state problem), then the matching of the wave function outside the potential $u(r) = Ae^{-\alpha r}$ to the solution inside $B \sin \kappa r$ at the boundary gives

$$\kappa \cot \kappa a = -\alpha \quad (19-63)$$

For $k \ll \kappa$, we have

$$\left(\frac{\tan \kappa a}{\kappa a} \right)_{\text{scatt}} \cong \left(\frac{\tan \kappa a}{\kappa a} \right)_{\text{deuteron}} = -\frac{1}{a\alpha} \quad (19-64)$$

Thus

$$\sigma \cong 4\pi a^2 \left(1 + \frac{1}{a\alpha}\right)^2 \cong \frac{4\pi}{\alpha^2} (1 + 2a\alpha) \quad (19-65)$$

Thus making the low energy approximation expressed by (19-64) allows us to bypass the problem of determining the potential and *then* calculating the cross section. The approximation only works when the binding energy is small. The quantity $1/\alpha$ is the distance over which the deuteron wave function spills over, and this is always much larger than the range of the potential a for a loosely bound system. It is $1/\alpha$ and not the range of the potential that determines the scattering cross section at low energies.

Spin-Dependent Scattering

In the 1930s there was great interest in the form of the neutron–proton potential, since it was hoped that this would give some fundamental clues concerning the nuclear forces in general. Rudimentary experiments at low energies were fitted with a variety of potentials. It became evident after a while that almost any reasonably shaped potential would work, provided that one chose the appropriate depth and range. It was shown in 1947 by Schwinger (and subsequently derived by Bethe in a simpler manner) that at low energies it is always a good approximation to write

$$k \cot \delta = -\frac{1}{A} + \frac{1}{2} r_0 k^2 \quad (19-66)$$

where A is called the scattering length, and r_0 is the effective range. The cross section at threshold determines the scattering length

$$\sigma \cong 4\pi A^2 \quad (19-67)$$

and the energy dependence determines the effective range. The relation between these parameters and the parameters describing the potential vary with the shape, but a two-parameter fit to the data is always possible. This *effective range formula* shows that if we want to probe the shape of the potential, we must go to higher energies.

The binding energy of the deuteron is 2.23 MeV. Thus, remembering that in our discussion m is the reduced mass—that is, $m \cong M_p/2$ —

$$\begin{aligned} \frac{1}{\alpha} &= \sqrt{\frac{\hbar^2}{2mE}} = \frac{\hbar c}{\sqrt{M_p c^2 E}} = \frac{\hbar}{M_p c} \sqrt{\frac{M_p c^2}{E}} \\ &= \frac{1.055 \times 10^{-34} \text{ J} \cdot \text{s}}{(1.67 \times 10^{-27} \text{ kg})(3 \times 10^8 \text{ m/s})} \sqrt{\frac{938 \text{ MeV}}{2.23 \text{ MeV}}} = 4.3 \times 10^{-15} \text{ m} \end{aligned}$$

and

$$\frac{4\pi}{\alpha^2} = 2.3 \times 10^{-28} \text{ m}^2 = 2.3 \text{ barns}$$

A more accurate determination leads to the prediction that the cross section at threshold is four barns. The measurement, carried out with neutrons at thermal speeds, yields 21 barns!

The explanation of this disagreement came with the realization that the spin of the neutron and the proton had not been taken into account. If the potential were spin independent, then all spin states would scatter the same way; that is, it would not matter

whether the spins of the particles are “up” or “down.” If the potential does depend on the spin, a possible form could be

$$V(r) = V_1(r) + \boldsymbol{\sigma}_p \cdot \boldsymbol{\sigma}_n V_2(r) \quad (19-68)$$

In this case spin is no longer a good quantum number, and the states must be classified by total angular momentum and total spin; that is, with $l = 0$, the four states divide up into a $^3S_1$ triplet of states, and a singlet $^1S_0$. These need not scatter the same way, so that there are really two phase shifts, δ_t for the triplet, and δ_s for the singlet. There are no triplet-singlet transitions, since the total angular momentum J must be the same in the initial and final states. The total cross section is weighted by the number of final states in each case (the cross section involves a *sum* over final states and is independent of the value of the z -component of the angular momentum), so that

$$\sigma = \frac{3}{4} \sigma_t + \frac{1}{4} \sigma_s \quad (19-69)$$

For spin independent forces, $\sigma = \sigma_t = \sigma_s$.

The deuteron is a $^3S_1$ state, so that the four barns are really predicted for σ_t . This implies that

$$\sigma_s = 4\sigma - 3\sigma_t = 72 \text{ barns} \quad (19-70)$$

Since we are at threshold, this implies that

$$|A_s| = \sqrt{\frac{72 \times 10^{-28}}{4\pi}} \cong 2.4 \times 10^{-14} \text{ m} \quad (19-71)$$

The earlier result implied that

$$|A_t| = \sqrt{\frac{4 \times 10^{-28}}{4\pi}} \cong 4.7 \times 10^{-13} \text{ m} \quad (19-72)$$

The question of the signs of A_t and A_s now arises. At threshold we have $k \cot \delta \approx k/\delta \approx -1/A$ so that $\delta_s = -A_s k$ and $\delta_t = -A_t k$. Thus, the asymptotic wave functions have the form

$$\sin(kr + \delta_{t,s}) \approx \sin k(r - A_{t,s}) \approx k(r - A_{t,s}) \quad (19-73)$$

The two possible cases are shown in Fig. 19-6. We know that for the triplet state the wave function turns over just before the edge of the well (since there is a bound state), so that it must correspond to the situation $A_t > 0$.

If A_s were positive, too, one would expect a singlet bound state, with very much weaker binding, since the internal wave function ties onto a much flatter asymptotic form. In fact, the binding energy would be 70 keV. Such a bound state was not found, suggesting that $A_s < 0$.

This choice of sign was actually confirmed by the scattering of neutrons off the H_2 molecule. As we know, the H_2 molecule can exist as ortho- H_2 , with the spins in a triplet state, and para- H_2 , with the two proton spins in a singlet state. For neutrons at very low energies, such that the wavelength is much larger than the proton-proton separation in the molecule, the scattering amplitude for neutron- H_2 , scattering is just the sum of the amplitudes for the individual scatterings. One may show that the amplitude off para- H_2 is different from the amplitude off ortho- H_2 and these separately involve linear combinations of A_s and A_t . The fact that $\sigma_{\text{para}} \cong 3.9$ barns, while $\sigma_{\text{ortho}} \cong 125$ barns can be explained in this way. The calculation is complicated by a number of effects that must be taken into

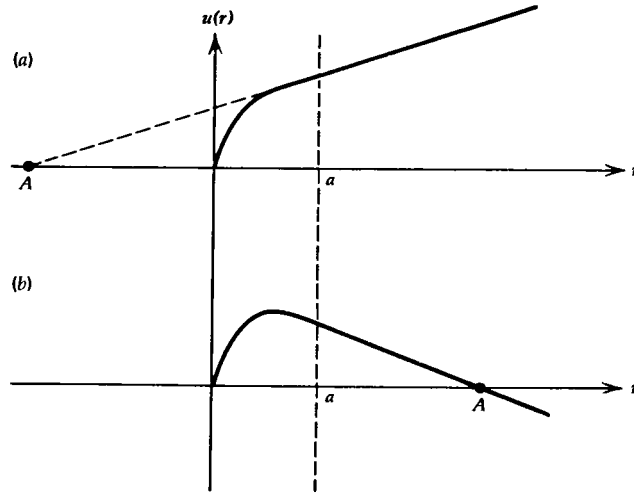


Figure 19-6 Sketch of the s -wave solution $u(r)$ near threshold. Outside the range radius $r = a$, the wave function has the form $C(r - A)$. [This is not in conflict with (19-73), which is an expansion of $\sin(kr + \delta)$. We could equally well have taken the form of $u(r)$ to be $(C/k) \sin(kr + \delta)$, since the normalization is arbitrary. It is, in fact, the interior wave function and the position of A that determine the slope of the line.] The sign of A depends on whether the interior wave function has or has not turned over cases (b) and (a), respectively. Since the wave function must turn over if there is a weakly bound state (so that it can match a slowly falling exponential) and since one does not expect the wave function inside the potential to be very sensitive to variations in E about zero, one expects that for a potential that has a bound state with E_B small, $A > 0$.

account, for example, that the effective mass of the proton in a molecule is different from that of a free proton, and that the molecules are not really at rest, but are moving with a distribution appropriate to the (low ~ 20 K) temperature. The large discrepancy between the two cross sections is not changed much by these corrections, and it can only be explained if A_s is indeed negative.

19-3 THE BORN APPROXIMATION

At higher energies many partial waves contribute to the scattering, and it is therefore preferable to avoid the angular momentum decomposition. A procedure that leads to a very useful approximation both when the potential is weak and when the energy is high is the Born approximation, in which we consider the scattering process as a transition, just like the transitions studied in Chapter 15. The difference is that here we consider the transitions

continuum $\rightarrow$ continuum

If we work in the center-of-mass system, we have effectively a one-particle problem, and this particle makes a transition from an initial state, described by the eigenfunction

$$\psi_i(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{p}_i \cdot \mathbf{r}/\hbar} \quad (19-74)$$

to the final state, described by

$$\psi_f(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{p}_f \cdot \mathbf{r}/\hbar} \quad (19-75)$$

where $\mathbf{p}_i$ and $\mathbf{p}_f$ are the initial and final momenta, respectively. The transition rate, following the Golden Rule (15-20) is given by

$$R_{i \rightarrow f} = \frac{2\pi}{\hbar} \int \frac{V d^3 \mathbf{p}_f}{(2\pi\hbar)^3} |M_{fi}|^2 \delta\left(\frac{p_f^2}{2m} - \frac{p_i^2}{2m}\right) \quad (19-76)$$

The delta function expresses energy conservation. If the particles that emerge have a different mass from those that enter, or if the target is excited, that delta function takes a somewhat different form. It will, however, always be of the form $\delta[(p_f^2/2m) - E]$ where E is the energy available for kinetic energy of the final particle. The matrix element M_{fi} is given by

$$\begin{aligned} M_{fi} &= \langle \psi_f | V | \psi_i \rangle = \int d^3 \mathbf{r} \frac{e^{-i\mathbf{p}_f \cdot \mathbf{r}/\hbar}}{\sqrt{V}} V(\mathbf{r}) \frac{e^{i\mathbf{p}_i \cdot \mathbf{r}/\hbar}}{\sqrt{V}} \\ &= \frac{1}{V} \int d^3 \mathbf{r} e^{-i\Delta \cdot \mathbf{r}} V(\mathbf{r}) \end{aligned} \quad (19-77)$$

where $\Delta = \frac{1}{\hbar} (\mathbf{p}_f - \mathbf{p}_i)$. We write the matrix element as

$$M_{fi} = \frac{1}{V} \tilde{V}(\Delta) \quad (19-78)$$

The integral in (19-76) can be rewritten in the form

$$\begin{aligned} R_{i \rightarrow f} &= \frac{2\pi}{\hbar} \int d\Omega \frac{V p_f^2 dp_f}{(2\pi\hbar)^3} \frac{1}{V^2} |\tilde{V}(\Delta)|^2 \delta\left(\frac{p_f^2}{2m} - E\right) \\ &= \frac{2\pi}{\hbar} \frac{1}{(2\pi\hbar)^3} \frac{1}{V} \int d\Omega p_f m \frac{p_f dp_f}{m} \delta\left(\frac{p_f^2}{2m} - E\right) |\tilde{V}(\Delta)|^2 \\ &= \frac{1}{4\pi^2 \hbar^4} \frac{1}{V} \int d\Omega p_f m |\tilde{V}(\Delta)|^2 \end{aligned} \quad (19-79)$$

To get the last line, we noted that $p_f dp_f/m = d(p_f^2/2m)$ and carried out the delta function integration. Thus, p_f must be evaluated at $p_f = (2mE)^{1/2}$, and we must not forget that m here is the reduced mass in the final state.

This expression has an undesirable dependence on the volume of the quantization box, but this is not really surprising. Our wave functions were normalized to one particle in the box V , so that the number of transitions should certainly go down as V increases. This difficulty arises because we are asking a question that does not correspond to an experiment. What one does is send a flux of incident particles at each other (in the center-of-mass frame; in the laboratory, one particle is stationary, of course). If we want a flux of one particle per m^2 per second, we must multiply the preceding by V divided by the volume of a cylinder with 1-m^2 base, and the relative velocity of the particles in the center-of-mass frame in the initial state. The number of transitions for unit flux is just the cross section. We therefore have

$$d\sigma = \frac{1}{4\pi^2 \hbar^4} \frac{1}{|v_{\text{rel}}|} d\Omega p_f m |\tilde{V}(\Delta)|^2 \quad (19-80)$$

Since in the center-of-mass frame the two incident particles are moving toward each other with equal and opposite momenta of magnitude p_i , their relative velocity is

$$|v_{\text{rel}}| = \frac{p_i}{m_1} + \frac{p_i}{m_2} = p_i \left(\frac{1}{m_1} + \frac{1}{m_2} \right) = \frac{p_i}{m_{\text{red}}^{(i)}} \quad (19-81)$$

if m_1 and m_2 are their masses. Thus, if the initial and final reduced masses and momenta are not the same, we have

$$\frac{d\sigma}{d\Omega} = \frac{1}{4\pi^2} \frac{p_f}{p_i} m_{\text{red}}^{(f)} m_{\text{red}}^{(i)} \left| \frac{1}{\hbar^2} \tilde{V}(\Delta) \right|^2 \quad (19-82)$$

When the initial and final particles are the same,

$$\frac{d\sigma}{d\Omega} = \frac{m_{\text{red}}^2}{4\pi^2} \left| \frac{1}{\hbar^2} \tilde{V}(\Delta) \right|^2 \quad (19-83)$$

When one particle is a great deal more massive than the other, $m_{\text{red}} \rightarrow m$, the mass of the lighter particle. When we compare the above with (19-15) we see that

$$f(\theta, \phi) = - \frac{m_{\text{red}}}{2\pi\hbar^2} \tilde{V}(\Delta) \quad (19-84)$$

Actually, to determine the sign, one must go through a more detailed comparison with the partial wave expansion. We will not bother to do this here.

As an illustration of the application of the Born approximation, we will calculate the cross section for the scattering of a particle of mass m and charge Z_1 by a Coulomb potential of charge Z_2 . The source of the Coulomb field is taken to be infinitely massive, so that the mass in (19-83) is the mass of the incident particle. For generality (and, as we will see, for technical reasons) we take the Coulomb field to be screened, so that

$$V(\mathbf{r}) = \frac{Z_1 Z_2 e^2}{4\pi\epsilon_0} \frac{e^{-r/a}}{r} \quad (19-85)$$

where a is the screening radius. We thus need to evaluate

$$\tilde{V}(\Delta) = \frac{Z_1 Z_2 e^2}{4\pi\epsilon_0} \int d^3\mathbf{r} e^{i\Delta\cdot\mathbf{r}} \frac{e^{-r/a}}{r} \quad (19-86)$$

We choose the direction of Δ as z -axis, and then get

$$\begin{aligned} \int d^3\mathbf{r} e^{-i\Delta\cdot\mathbf{r}} \frac{e^{-r/a}}{r} &= \int_0^{2\pi} d\phi \int_0^\pi \sin\theta d\theta \int_0^\infty r^2 dr e^{-i\Delta r \cos\theta} \frac{e^{-r/a}}{r} \\ &= 2\pi \int_0^\infty r dr e^{-r/a} \int_{-1}^1 d(\cos\theta) e^{-i\Delta r \cos\theta} \\ &= \frac{2\pi}{i\Delta} \int_0^\infty dr e^{-r/a} (e^{i\Delta r} - e^{-i\Delta r}) \\ &= \frac{2\pi}{i\Delta} \left(\frac{1}{(1/a) - i\Delta} - \frac{1}{(1/a) + i\Delta} \right) = \frac{4\pi}{(1/a^2) + \Delta^2} \end{aligned} \quad (19-87)$$

Now

$$\Delta^2 = \frac{1}{\hbar^2} (\mathbf{p}_f - \mathbf{p}_i)^2 = \frac{1}{\hbar^2} (2p^2 - 2\mathbf{p}_f \cdot \mathbf{p}_i) = \frac{2p^2}{\hbar^2} (1 - \cos\theta) \quad (19-88)$$

so that the cross section becomes

$$\begin{aligned}
 \frac{d\sigma}{d\Omega} &= \frac{m^2}{4\pi^2\hbar^4} \left(\frac{Z_1 Z_2 e^2}{4\pi\epsilon_0} \right)^2 \frac{16\pi^2}{[(2p^2/\hbar^2)(1 - \cos\theta) + (1/a^2)^2]^2} \\
 &= \left(\frac{2m \frac{Z_1 Z_2 e^2}{4\pi\epsilon_0}}{4p^2 \sin^2(\theta/2) + (\hbar^2/a^2)} \right)^2 \\
 &= \left(\frac{\frac{Z_1 Z_2 e^2}{4\pi\epsilon_0}}{4E \sin^2(\theta/2) + (\hbar^2/2ma^2)} \right)^2
 \end{aligned} \tag{19-89}$$

In the last line we replaced $p^2/2m$ by E , and we used $\frac{1}{2}(1 - \cos\theta) = \sin^2(\theta/2)$. The angle θ defined in (19-88) is the center-of-mass scattering angle. In the absence of screening ($a \rightarrow \infty$) this reduces to the well-known Rutherford formula. There is no $\hbar$ in it, and it is the same as the classical formula. Had we left out the screening factor in (19-86) we would have had an ill-defined integral. One often evaluates ambiguous integrals with the aid of such convergence factors.

The Born approximation has its limitations. For example, we found that $\tilde{V}(\Delta)$ was purely real so that $f(\theta)$ is also real in this approximation. This implies, by the optical theorem, that the cross section is zero. In fact, the Born approximation is only good when either (a) the potential is weak, so that the cross section is of second order in a small parameter; this would make the use of it consistent with the optical theorem, or (b) at high energies for potentials such that the cross section goes to zero. This is true for most smooth potentials. It is not true for real particles; there it seems that the cross sections stay constant at very high energies, and one cannot expect the Born approximation to serve as more than a guide of the behavior of the scattering amplitude.

As a final comment, we observe that if the potential V has a spin dependence, then (19-77) is trivially modified by the requirement that the initial and final states be described by their spin wave functions, in addition to the spatial wave functions. Thus, for example, if the neutron-proton potential has the form

$$V(r) = V_1(r) + \boldsymbol{\sigma}_P \cdot \boldsymbol{\sigma}_N V_2(r)$$

the Born approximation reads

$$M_{fi} = \frac{1}{V} \int d^3r e^{-i\Delta \cdot r} \xi_f^+ V(r) \xi_i$$

where ξ_i and ξ_f represent the initial and final spin states of the neutron-proton system.

19-4 SCATTERING OF IDENTICAL PARTICLES

When two identical particles scatter, there is no way of distinguishing a deflection of a particle through an angle θ and a deflection of $\pi - \theta$ in the center-of-mass frame, since momentum conservation demands that if one of the particles scatters through θ , the other goes in the direction $\pi - \theta$ (Fig. 19-7). Classically, too, the cross section for scattering is affected by the identity of the particles, since the number of counts at a certain counter will be the sum of the counts due to the two particles. Thus

$$\sigma_d(\theta) = \sigma(\theta) + \sigma(\pi - \theta) \tag{19-90}$$

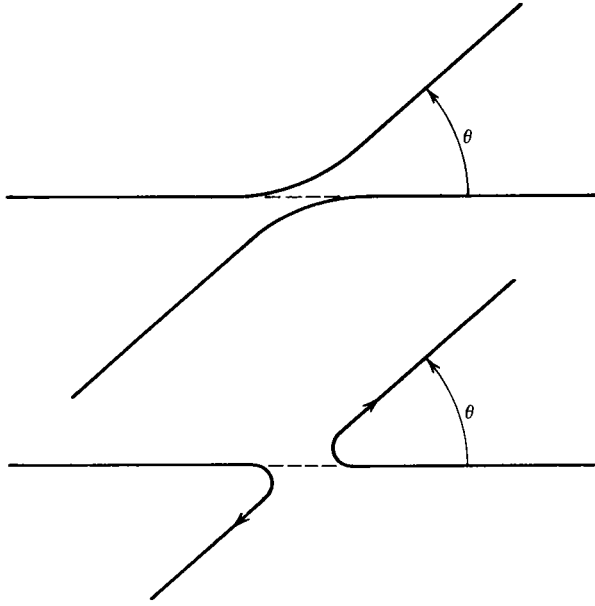


Figure 19-7 Asymptotic directions in the scattering of two identical particles through a center-of-mass angle θ .

In quantum mechanics there is no way of distinguishing the two final states, so that the two *amplitudes* $f(\theta)$ and $f(\pi - \theta)$ can interfere. Thus the cross section for the scattering of two identical spin zero (boson) particles—for example, α -particles—is

$$\frac{d\sigma}{d\Omega} = |f(\theta) + f(\pi - \theta)|^2 \quad (19-91)$$

This differs from the classical result by the interference term

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 + |f(\pi - \theta)|^2 + [f^*(\theta)f(\pi - \theta) + f(\theta)f^*(\pi - \theta)] \quad (19-92)$$

and it leads to an enhancement at $\pi/2$ —for example,

$$\left(\frac{d\sigma}{d\Omega}\right)_{\pi/2} = 4 \left|f\left(\frac{\pi}{2}\right)\right|^2 \quad (19-93)$$

compared to the result that would be obtained without interference:

$$\left(\frac{d\sigma}{d\Omega}\right)_{\pi/2} = 2 \left|f\left(\frac{\pi}{2}\right)\right|^2 \quad (19-94)$$

When the scattering of two spin $1/2$ particles is considered—for example, proton–proton scattering or electron–electron scattering—then the amplitude should reflect the basic antisymmetry of the total wave function under the interchange of the two particles. If the two particles are in a spin singlet state, then the spatial wave function is symmetric, and

$$\frac{d\sigma_s}{d\Omega} = |f(\theta) + f(\pi - \theta)|^2 \quad (19-95)$$

If the two particles are in a spin triplet state, then the spatial wave function is antisymmetric, and

$$\frac{d\sigma_t}{d\Omega} = |f(\theta) - f(\pi - \theta)|^2 \quad (19-96)$$

In the scattering of two unpolarized protons, all spin states are equally likely, and thus the probability of finding the two protons in a triplet state is three times as large as finding them in a singlet state, so that

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \frac{3}{4} \frac{d\sigma_t}{d\Omega} + \frac{1}{4} \frac{d\sigma_s}{d\Omega} \\ &= \frac{3}{4} |f(\theta) - f(\pi - \theta)|^2 + \frac{1}{4} |f(\theta) + f(\pi - \theta)|^2 \\ &= |f(\theta)|^2 + |f(\pi - \theta)|^2 - \frac{1}{2} [f(\theta)f^*(\pi - \theta) + f^*(\theta)f(\pi - \theta)]\end{aligned}\quad (19-97)$$

For proton-proton scattering as well as for α - α scattering, the basic amplitude $f(\theta)$ is the sum of a nuclear term (if the energies are not too low) and a Coulomb term. Whether the identical particles are bosons or fermions, there is a symmetry under the interchange $\theta \rightarrow \pi - \theta$.

Scattering by Atoms on a Lattice

Symmetry considerations also play a role in the scattering of particles by a crystal lattice. If we ignore spin, so that we do not have to worry whether the electron does or does not flip its spin ("up" $\rightarrow$ "down" or vice versa), then at low energies, the scattering amplitude $f(\theta)$ is independent of angle (*S*-wave scattering), and the solution of the Schrödinger equation by a single atom located at the lattice point $\mathbf{a}_i$ has the asymptotic form

$$\psi(\mathbf{r}) \sim e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{a}_i)} + f \frac{e^{ik|\mathbf{r} - \mathbf{a}_i|}}{|\mathbf{r} - \mathbf{a}_i|} \quad (19-98)$$

Now

$$\begin{aligned}k|\mathbf{r} - \mathbf{a}_i| &= k(\mathbf{r}^2 - 2\mathbf{r} \cdot \mathbf{a}_i + a_i^2)^{1/2} \\ &\cong kr \left(1 - \frac{2\mathbf{r} \cdot \mathbf{a}_i}{r^2}\right)^{1/2} \\ &\cong kr - k\hat{\mathbf{r}} \cdot \mathbf{a}_i\end{aligned}\quad (19-99)$$

and since $k\hat{\mathbf{r}}$ is a vector of magnitude k and it points in the direction $\mathbf{r}$, the point of observation, it is the final momentum $\mathbf{k}'$. If we divide out the phase factor $e^{-i\mathbf{k}' \cdot \mathbf{a}_i}$, the wave function has the asymptotic form

$$\psi \sim e^{i\mathbf{k} \cdot \mathbf{r}} + f e^{-i\mathbf{k}' \cdot \mathbf{a}_i} e^{i\mathbf{k} \cdot \mathbf{a}_i} \frac{e^{ikr}}{r} + O\left(\frac{1}{r^2}\right) \quad (19-100)$$

so that the scattering amplitude is

$$f(\theta) = f e^{-i\Delta \cdot \mathbf{a}_i}, \quad \Delta = \mathbf{k}' - \mathbf{k} \quad (19-101)$$

The total amplitude is the sum of all individual scattering amplitudes when we have a situation in which we cannot tell which atom in the crystal did the scattering. This is indeed the case for elastic low-energy scattering when recoil is not observed and spins are not measured. Thus, for the *coherent* process we have

$$\frac{d\sigma}{d\Omega} = \left| f \sum_{\text{atoms}} e^{-i\Delta \cdot \mathbf{a}_i} \right|^2 \quad (19-102)$$

If we have a simple cubic array of lattice points, such that

$$\mathbf{a}_i = a(n_x \hat{\mathbf{i}}_x + n_y \hat{\mathbf{i}}_y + n_z \hat{\mathbf{i}}_z) \quad -N \leq n_x, n_y, n_z \leq N \quad (19-103)$$

(spacings are integral multiples of a in all directions), then

$$\sum e^{-i\Delta \cdot \mathbf{a}_i} = \sum_{n_x=-N}^N \sum_{n_y=-N}^N \sum_{n_z=-N}^N e^{-ia\Delta_x n_x} e^{-ia\Delta_y n_y} e^{-ia\Delta_z n_z}$$

We use

$$\begin{aligned}
 \sum_{n=-N}^N e^{i\alpha n} &= e^{-i\alpha N} (1 + e^{i\alpha} + e^{2i\alpha} + \dots + e^{2i\alpha N}) \\
 &= e^{i\alpha N} \frac{e^{i\alpha(2N+1)} - 1}{e^{i\alpha} - 1} = \frac{e^{i\alpha(N+1)} - e^{-i\alpha N}}{e^{i\alpha} - 1} \\
 &= \frac{e^{i\alpha(N+1/2)} - e^{-i\alpha(N+1/2)}}{e^{i\alpha/2} - e^{-i\alpha/2}} = \frac{\sin \alpha (N + \frac{1}{2})}{\sin \alpha/2}
 \end{aligned} \tag{19-104}$$

to obtain the result

$$\frac{d\sigma}{d\Omega} = |f|^2 \frac{\sin^2 \alpha_x (N + \frac{1}{2})}{\sin^2 \alpha_x/2} \cdot \frac{\sin^2 \alpha_y (N + \frac{1}{2})}{\sin^2 \alpha_y/2} \cdot \frac{\sin^2 \alpha_z (N + \frac{1}{2})}{\sin^2 \alpha_z/2} \tag{19-105}$$

where

$$\alpha_x = a \Delta_x - 2\pi\nu_x \quad (\nu_x = \text{integer}), \text{ etc.} \tag{19-106}$$

We can make the generalization exhibited earlier, since a change $\alpha \rightarrow \alpha - 2\pi\nu$, with ν an integer, does not change (19-105). The expression (19-105) is not very transparent. However, when N is large, each of the factors becomes very strongly peaked when $\alpha_x, \dots$ are near zero. In fact, using

$$\frac{\sin^2 Nu}{u^2/4} \rightarrow 4\pi N \delta(u) \tag{19-107}$$

a formula easily derived from (15-14) by a simple change of variables, we get

$$\frac{d\sigma}{d\Omega} = |f|^2 (2\pi)^3 (2N)^3 \delta(a\Delta - 2\pi\nu) \tag{19-108}$$

Now the total number of atoms is $(2N)^3$, and hence the cross section per atom is

$$\frac{d\sigma}{d\Omega} = |f|^2 \frac{(2\pi)^3}{a^3} \delta\left(\mathbf{k}' - \mathbf{k} - \frac{2\pi\nu}{a}\right) \tag{19-109}$$

Thus the differential cross section is very small, except in the directions given by

$$\mathbf{k}' - \mathbf{k} = \frac{2\pi}{a} \boldsymbol{\nu} \tag{19-110}$$

where it is strongly peaked. The conditions above are called the *Bragg conditions*, and the integers ν_x, ν_y, ν_z are called the *Miller indices of the Bragg planes*.

The relations just derived can be generalized to more complicated crystals. They are used to study crystal structure, using neutrons or X rays as incident particles, or using a known crystal to study X rays that are emitted in atomic transitions involving energetic photons.

19-5 AN INELASTIC PROCESS: THE PHOTOELECTRIC EFFECT

An interesting example of an inelastic scattering process is the photoelectric effect. It is of interest in that it plays an important role in the absorption of radiation in matter in much of the nonrelativistic energy region. The basic process is

$$\gamma + (\text{atom}) \rightarrow (\text{atom})^+ + e^-$$

Following the golden rule [cf. (15-23)] the transition rate for this process is

$$\begin{aligned}
 R &= \frac{2\pi}{\hbar} \int \frac{V d^3 p}{(2\pi\hbar)^3} |M_{fi}|^2 \delta\left(\hbar\omega - E_B - \frac{p_e^2}{2m}\right) \\
 &= \frac{2\pi}{\hbar} \int \frac{V d\Omega}{(2\pi\hbar)^3} \int m p_e d\left(\frac{p_e^2}{2m}\right) |M_{fi}|^2 \delta\left(\hbar\omega - E_B - \frac{p_e^2}{2m}\right) \\
 &= \frac{2\pi V}{\hbar} \int d\Omega \frac{m p_e}{(2\pi\hbar)^3} |M_{fi}|^2
 \end{aligned} \quad (19-111)$$

In this expression m is the electron mass, the delta function represents energy conservation, E_B is the magnitude of the binding energy of the electron in the atom, and in the last line, p_e is evaluated at the vanishing of the argument of the delta function.

The matrix element is given by

$$\frac{e}{m} \sqrt{\frac{\hbar}{2\varepsilon_0\omega V}} \int d^3 r \psi_f^*(\mathbf{r}) \boldsymbol{\varepsilon} \cdot \mathbf{p} e^{i\mathbf{k}\cdot\mathbf{r}} \psi_i(\mathbf{r}) \quad (19-112)$$

The vector potential is normalized, as in (17-15), to one photon in the volume V , and ψ_i, ψ_f are the initial and final state (normalized) wave functions for the electron. We assume that the electron is initially in the ground state of a hydrogenlike atom, so that

$$\psi_i(\mathbf{r}) = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0} \quad (19-113)$$

The final state wave function should be taken to be a solution of the Schrödinger equation with a Coulomb potential with $E > 0$. We did not discuss these solutions when we studied the hydrogen atom. They can be written in closed form, but they are quite complicated, and make integrals such as appear in (19-112) very difficult. If the photon energy is much larger than the ionization energy ($13.6Z^2$ eV), then the residual interaction of the outgoing electron with the ion that it leaves behind is less important. We may therefore replace the final electron wave function by a free-particle wave function. Since we started with only one atom in our volume V , there will only be one free electron in that volume. This means that the *normalized* final state wave function is

$$\psi_f(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{p}_e \cdot \mathbf{r} / \hbar} \quad (19-114)$$

The V that appears in the phase space factor $V d^3 p / (2\pi\hbar)^3$ has the same origin, so that the two factors are not independent. The matrix element is simplified, since the final state is an eigenstate of the momentum so that

$$\langle f | \boldsymbol{\varepsilon} \cdot \mathbf{p}_{\text{op}} e^{i\mathbf{k}\cdot\mathbf{r}} | i \rangle = \boldsymbol{\varepsilon} \cdot \mathbf{p}_e \langle f | e^{i\mathbf{k}\cdot\mathbf{r}} | i \rangle$$

The square of the matrix element then becomes

$$|M_{fi}|^2 = \left(\frac{e}{m}\right)^2 \frac{\hbar}{2\varepsilon_0\omega V} \frac{1}{V} \frac{1}{\pi} \left(\frac{Z}{a_0}\right)^3 (\boldsymbol{\varepsilon} \cdot \mathbf{p}_e)^2 \left| \int d^3 r e^{i(\mathbf{k} - \mathbf{p}_e/\hbar) \cdot \mathbf{r}} e^{-Zr/a_0} \right|^2 \quad (19-115)$$

As in our discussion of scattering, we are interested in the cross section, and we multiply the rate by V and divide by the relative velocity of the particles in the initial state. Since one of them is a photon, that velocity is always c . We thus end up with

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{\hbar} \frac{V m p_e}{(2\pi\hbar)^3} \frac{V}{c} \left(\frac{e}{m}\right)^2 \frac{\hbar}{2\varepsilon_0\omega V} \frac{1}{V} \frac{1}{\pi} \left(\frac{Z}{a_0}\right)^3 (\boldsymbol{\varepsilon} \cdot \mathbf{p}_e)^2 \left| \int d^3 r e^{i(\mathbf{k} - \mathbf{p}_e/\hbar) \cdot \mathbf{r}} e^{-Zr/a_0} \right|^2 \quad (19-116)$$

In this expression $d\Omega$ is the solid angle into which $\mathbf{p}_e$ points. The integral over all electron direction yields the total cross section σ for the photoelectric effect. If the target atoms are distributed with a density of N atoms/m³, then in a slab of material of area A and thickness dx , there are $Nadx$ target atoms. Each atom presents an effective area σ for the reaction under consideration, so that the total area presented to the beam is $NA\sigma dx$. If there are n incident particles in the bombarding beam, then the number of particles dn that interact in the thickness dx of the target is given by the formula

$$\frac{dn}{n} = -\frac{NA\sigma dx}{A} = -N\sigma dx \quad (19-117)$$

The minus sign is there because particles are *removed* from the beam. Integration gives

$$n(x) = n_0 e^{-N\sigma x} \quad (19-118)$$

where n_0 is the number of incident particles and $n(x)$ is the number of particles left in the beam after it traverses a thickness x of target. The quantity $\lambda = 1/N\sigma$ has the dimensions of a length, and it is called the *mean free path*.

The remaining task is to evaluate the integral in the expression for $d\sigma/d\Omega$. The integral is of the form

$$\int d^3r e^{i\mathbf{q}\cdot\mathbf{r}} e^{-\beta r} = -\frac{d}{d\beta} \int d^3r e^{i\mathbf{q}\cdot\mathbf{r}} \frac{e^{-\beta r}}{r} \quad (19-119)$$

The reason for writing the integral in this form is that the second integral was worked out in (19-87), with the result

$$\int d^3r e^{i\mathbf{q}\cdot\mathbf{r}} \frac{e^{-\beta r}}{r} = \frac{4\pi}{\beta^2 + q^2}$$

Thus

$$\int d^3r e^{i\mathbf{q}\cdot\mathbf{r}} e^{-\beta r} = \frac{8\pi\beta}{(\beta^2 + q^2)^2}$$

and

$$\int d^3r e^{i(\mathbf{k} - \mathbf{p}_e/\hbar)\cdot\mathbf{r}} e^{-Z\hbar/a_0} = \frac{8\pi(Z/a_0)}{[(\mathbf{k} - \mathbf{p}_e/\hbar)^2 + (Z/a_0)^2]^2} \quad (19-120)$$

We can now calculate the differential cross section. After some judicious combination of factors, we get

$$\frac{d\sigma}{d\Omega} = 32Z^5 a_0^2 \left(\frac{p_e c}{\hbar\omega} \right) \left(\frac{\mathbf{\epsilon} \cdot \mathbf{p}_e}{mc} \right)^2 \frac{1}{(Z^2 + a_0^2 \Delta^2)^4} \quad (19-121)$$

where

$$\Delta^2 = \left(\frac{\hbar\mathbf{k} - \mathbf{p}_e}{\hbar} \right)^2 = \left(\frac{\mathbf{p}_\gamma - \mathbf{p}_e}{\hbar} \right)^2 \quad (19-122)$$

Since the electron and photon energies are related by

$$\hbar\omega = E_B + \frac{p_e^2}{2m_e} \quad (19-123)$$

we see that for energies quite a bit above the ionization energy, $\hbar\omega \equiv p_e^2/2m_e$. Furthermore,

$$\begin{aligned}\Delta^2 &= \frac{1}{\hbar^2} (\mathbf{p}_\gamma - \mathbf{p}_e)^2 = \frac{1}{\hbar^2} \left[\left(\frac{\hbar\omega}{c} \right)^2 - 2 \frac{\hbar\omega}{c} p_e \cos \theta + p_e^2 \right] \\ &\cong \frac{1}{\hbar^2} \left[p_e^2 - \frac{p_e^3}{m_e c} \cos \theta \right] \\ &\cong \frac{p_e^2}{\hbar^2} \left[1 - \frac{v_e}{c} \cos \theta \right]\end{aligned}\quad (19-124)$$

for nonrelativistic⁵ electrons, $p_e \ll mc$. In the above θ is the angle between the photon and the electron direction. If we choose the photon direction to define the z -axis, and the two photon polarization directions $\mathbf{e}^{(1)}$ and $\mathbf{e}^{(2)}$ to point in the x - and y -directions, respectively, then, writing

$$\hat{\mathbf{p}}_e = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) \quad (19-125)$$

we have $(\hat{\mathbf{p}}_e \cdot \mathbf{e}^{(1)})^2 = \sin^2 \theta \cos^2 \varphi$ and $(\hat{\mathbf{p}}_e \cdot \mathbf{e}^{(2)})^2 = \sin^2 \theta \sin^2 \varphi$, so that the average⁶ of $(\mathbf{e} \cdot \mathbf{p}_e)^2$ is $(1/2)$ the sum of these; that is, $\frac{1}{2} \sin^2 \theta$. We now substitute all of these into (19-121) and note that *light* elements, $\hbar\omega \gg E_B$, is equivalent to

$$E_e \gg \frac{1}{2} m_e c^2 (Z\alpha)^2$$

Putting all of this together we end up with

$$\frac{d\sigma}{d\Omega} = 2\sqrt{2} Z^7 \alpha^8 \left(\frac{a_0}{Z} \right)^2 \left(\frac{E_e}{m_e c^2} \right)^{-7/2} \frac{\sin^2 \theta}{\left(1 - \frac{v_e}{c} \cos \theta \right)^4} \quad (19-126)$$

Let us discuss various aspects of this formula.

(1) First, the vague guess that, since atomic sizes tend to be of the order of 10^{-10} m, the cross sections should be of order 10^{-20} m² is wrong! It is true that the factor a_0^2 is of that magnitude, but it is multiplied by $(1/137)^8$, which is dimensionless, but hardly negligible! We should try to understand how one could be so wrong in order to have some guidance on what one must be careful about in making estimates. If we ignore the last angular factor, which we will discuss later, we see that we may, with the help of

$$E = \frac{1}{2} m_e v_e^2$$

write the factor in front as

$$\begin{aligned}2\sqrt{2} a_0^2 Z^5 \alpha^8 \left(\frac{m_e c^2}{E} \right)^{7/2} &= 32 a_0^2 Z^5 \alpha^8 \left(\frac{c}{v_e} \right)^7 \\ &= 32 \left(\frac{a_0}{Z} \right)^2 \alpha \left(\frac{\alpha Z c}{v_e} \right)^7\end{aligned}\quad (19-127)$$

This is a more useful form. It shows, first of all, the presence of a single factor α , which should always be present when a single photon is emitted or absorbed. The coupling of

⁵For relativistic electrons one should really use the Dirac equation to describe the process. As far as the stopping of radiation in matter, effects other than the photoelectric one are important when E_γ gets up to 1 MeV.

⁶By averaging we are calculating the photoeffect cross section for unpolarized photons.

the vector potential to a charge is proportional to the charge e , and the square of this will lead to the α . The factor $(a_0/Z)^2$ is a better measure of the area of the atom than a_0^2 , since we are considering a hydrogenlike atom of charge Z . What remains is a rather high power of the ratio of the “orbital” velocity of the electron in the atom to the velocity of the outgoing free electron.

It is the ratio $(\alpha Zc/v_e)$ [rather than just (c/v_e) , which is also dimensionless] that appears, because the matrix element involves the overlap between the free electron wave function and the bound electron wave function; that is, the square matrix element is related to the probability that a measurement of the momentum of the bound electron yields p_e . The functional dependence $f(\alpha Zc/v_e)$, in this case the eighth power,⁷ cannot be guessed at on general qualitative grounds. For example, if the electron wave function were gaussian [$\psi_i(\mathbf{r}) \propto e^{-r^2/a^2}$], the falloff with increasing velocity would be much faster than the eighth power. The reason why a guess is hard to make is that the momentum distribution of the electron is localized in a region of spread

$$\Delta p \sim \frac{\hbar}{a_0/Z} \sim \frac{\hbar Z}{\hbar/m_e c \alpha} \sim Z \alpha m_e c \quad (19-128)$$

and for $p_e \gg Z \alpha m_e c$ one is far out in the tail of the momentum distribution. This, again by the uncertainty relation, depends on the small r -distribution of the wave function, and depends sensitively on the state, in particular on the angular momentum. This does make photodisintegration in nuclear physics a very useful tool.

(2) The angular distribution of $d\sigma/d\Omega$ is given by

$$F(\theta) = \frac{\sin^2 \theta}{[1 - (v_e/c) \cos \theta]^4} \quad (19-129)$$

We note, first of all, that the cross section vanishes in the forward direction. This is a consequence of the fact that photons are transversely polarized. The matrix element is proportional to $\mathbf{p}_e \cdot \boldsymbol{\varepsilon}$, and when $\mathbf{p}_e$ is parallel to the photon momentum, this factor vanishes. The factor in the denominator has, because of the fourth power, a strong influence on the angular distribution. When v_e/c approaches unity, this becomes very dramatic, but even for moderate v_e/c there is significant peaking in the near forward direction, where the denominator is at its smallest. This corresponds to the minimum value of the momentum transfer between photon and electron $(\mathbf{p}_\gamma - \mathbf{p}_e)^2$.

PROBLEMS

1. Show that for a central potential $V(\mathbf{r}) = V(r)$, the matrix element M_{fi} in (19-77) can be written in the form

$$M_{fi} = \frac{1}{V} \frac{4\pi\hbar}{\Delta} \int_0^\infty r dr V(r) \sin r\Delta$$

Note that this is an even function of Δ —that is, a function of

$$\Delta^2 = (\mathbf{p}_f - \mathbf{p}_i)^2/\hbar^2$$

2. Consider a potential of the form

$$V(r) = V_0 e^{-r^2/a^2}$$

⁷There is a factor p_e in the phase space so that the matrix element squared gives an eighth power of $(\alpha Zc/v_e)$.

Calculate, using the Born approximation, the differential cross section $d\sigma/d\Omega$ as a function of the center-of-mass scattering angle θ . Compare your result with the differential cross section for a Yukawa potential

$$V(r) = V_0 b \frac{e^{-r/b}}{r}$$

[Already done in (19-85)–(19-89)]. To make the comparison, adjust the parameters in the two cases so that the two differential cross sections and their slopes are the same in the forward direction at $\Delta = 0$. It might be convenient to pick some definite numerical values for V_0 , b , a , and b to depict this graphically. Can you give a qualitative argument explaining the large difference between the predictions for large momentum transfers?

3. Consider the potential

$$V(r) = V_0 a \frac{e^{-r/a}}{r}$$

If the range parameter is $a = 1.2 \text{ fm} = 1.2 \times 10^{-15} \text{ m}$ and $V_0 = 100 \text{ MeV}$ in magnitude, what is the total cross section for proton–proton scattering at 100-MeV center-of-mass energy, calculated in Born approximation? Ignore Coulomb scattering, but not identity of two protons. [Note: It is useful to use the relation

$$\hbar^2 \Delta^2 = (\mathbf{p}_f - \mathbf{p}_i)^2 = 2p^2(1 - \cos \theta)$$

to write

$$d\Omega = 2\pi d(\cos \theta) = \frac{\hbar^2 \pi}{p^2} d(\Delta^2).]$$

4. Suppose the scattering amplitude for neutron–proton scattering is given by the form

$$f(\theta) = \xi_f^\dagger (A + B\sigma_P \cdot \sigma_N) \xi_i$$

where ξ_i and ξ_f are the initial and final spin states of the neutron–proton system. The possible states are

$$\begin{array}{ll} \xi_i = \chi_{\uparrow}^{(P)} \chi_{\uparrow}^{(N)} & \xi_f = \chi_{\uparrow}^{(P)} \chi_{\uparrow}^{(N)} \\ \chi_{\uparrow}^{(P)} \chi_{\downarrow}^{(N)} & \chi_{\uparrow}^{(P)} \chi_{\downarrow}^{(N)} \\ \chi_{\downarrow}^{(P)} \chi_{\uparrow}^{(N)} & \chi_{\downarrow}^{(P)} \chi_{\uparrow}^{(N)} \\ \chi_{\downarrow}^{(P)} \chi_{\downarrow}^{(N)} & \chi_{\downarrow}^{(P)} \chi_{\downarrow}^{(N)} \end{array}$$

Use

$$\sigma_P \cdot \sigma_N = \sigma_z^{(P)} \sigma_z^{(N)} + 2(\sigma_+^{(P)} \sigma_-^{(N)} + \sigma_-^{(P)} \sigma_+^{(N)})$$

where

$$\sigma_+ = \frac{\sigma_x + i\sigma_y}{2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \sigma_- = \frac{\sigma_x - i\sigma_y}{2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

in the representation in which $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\chi_{\uparrow} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\chi_{\downarrow} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to calculate all 16 scattering amplitudes. Make a table of your results and also tabulate the cross sections.

5. If any one of the spin states (e.g., initial proton, or initial neutron) is not measured, the cross section is the sum over the unmeasured spin states. Suppose both the initial and final proton spins are not measured. Write down expressions for the cross sections for the final neutron “up” and the final neutron “down,” given that the initial neutron state is “up.” What is the polarization P , defined by

$$P = \frac{\sigma_{\uparrow} - \sigma_{\downarrow}}{\sigma_{\uparrow} + \sigma_{\downarrow}}$$

where $\sigma_{\uparrow}$ is the cross section with the final neutron up and so on?

6. Use the table computed in Problem 4 to calculate the cross sections for triple $\rightarrow$ triplet and singlet $\rightarrow$ singlet scattering, respectively. Show the triplet $\rightarrow$ singlet scattering vanishes. Check your results by observing that since (in units of $\hbar$)

$$\frac{1}{2}\boldsymbol{\sigma}_P + \frac{1}{2}\boldsymbol{\sigma}_N = \mathbf{S}$$

one has

$$\begin{aligned}\boldsymbol{\sigma}_P \cdot \boldsymbol{\sigma}_N &= 2\mathbf{S}^2 - 3 \\ &= 1 \quad \text{when acting on triplet state} \\ &= -3 \quad \text{when acting on singlet state}\end{aligned}$$

Note that the amplitude is independent of m_S so that m_S must be the same in the initial and final spin states. There are three states in the triplet, all contributing an equal amount to the cross section, and only one to the singlet cross section. [Caution: In calculating amplitudes such as

$$\frac{1}{\sqrt{2}}(\chi_{\uparrow}^{(P)}\chi_{\downarrow}^{(N)}) - \chi_{\downarrow}^{(P)}\chi_{\uparrow}^{(N)})(A + B\boldsymbol{\sigma}_P \cdot \boldsymbol{\sigma}_N)\frac{1}{\sqrt{2}}(\chi_{\uparrow}^{(P)}\chi_{\downarrow}^{(N)}) - \chi_{\downarrow}^{(P)}\chi_{\uparrow}^{(N)}$$

the amplitudes are added for the four terms before squaring. Can you explain why?]

7. Consider the integral

$$I(kr) = \int_0^\pi d\theta \sin \theta g(\cos \theta) e^{-ikr \cos \theta} = \int_{-1}^1 du g(u) e^{-ikru}$$

where $g(\cos \theta)$ is strongly localized about $\theta = \theta_0$, and is infinitely differentiable. An example of such a function is

$$g = e^{-\alpha^2(\cos \theta - \cos \theta_0)^2}$$

with α large. We may thus assume that $g(u)$ and all of its derivatives at $u = \pm 1$ vanish. In that case, show that $I(kr)$ vanishes faster than any power of kr as $kr \rightarrow \infty$. (Hint: Write $e^{-ikru} = i/kr d/du e^{-ikru}$ and integrate by parts repeatedly.)

8. Calculate the cross section for the process

$$\gamma + \text{deuteron} \rightarrow N + P$$

The procedure is the same as that for the photoelectric effect. In the calculation of the matrix element, the final state wave function is again

$$\psi_f(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}$$

where $\mathbf{p}$ is the proton momentum. At low energies the wavelength of the radiation is much larger than the “size” of the deuteron, so that $e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1$. To calculate

$$\int d^3r e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar} \psi_i(r)$$

use

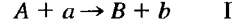
$$\begin{aligned}\psi_i(\mathbf{r}) &= \frac{N}{\sqrt{4\pi}} e^{-\alpha(r-r_0)} & r > r_0 \\ &= 0 & r < r_0\end{aligned}$$

properly normalized. For what energies would you expect the photon wavelength to be much larger than the range of the potential $r_0 \cong 1.2$ fm?

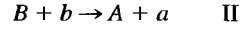
9. Suppose the electron were bound to the nucleus by a square well potential. Calculate the energy dependence of the cross section for the photoelectric effect. Assume that the photon energy is much larger

than the binding energy of the electron, and that the potential has a short range. (*Hint:* See Problem 8.)

10. The principle of detailed balance relates the matrix elements for the reactions



and



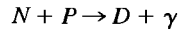
Thus

$$\sum |M_I|^2 = \sum |M_{II}|^2$$

where the sum is over both initial and final spin states. Taking into account that in the calculation of a rate or cross section one averages over the initial spin states and sums over the final spin states, show that for the rates

$$\frac{(2J_A + 1)(2J_a + 1)}{p_b^2(dp_b/dE_b)} \frac{dR_I}{d\Omega_b} = \frac{(2J_B + 1)(2J_b + 1)}{p_a^2(dp_a/dE_a)} \frac{dR_{II}}{d\Omega_a}$$

where J_a, J_A, J_b, J_B are the spins of the particles, p_b and p_a are the center-of-mass momenta of particles b and a (I and II must take place at the same total energy), E_b and E_a are the corresponding energies of the particles, and $d\Omega_b, d\Omega_a$ are the solid angles in which b and a are observed. Use this result to express the cross section for the radiative capture process



in terms of the cross section calculated in Problem 8. Note that the factor $(2J + 1)$ for photons is 2 since there are only two polarization states, and also the spin of the deuteron is 1.

Chapter 20

Entanglement and Its Implications

In the previous chapters we have exposed the reader to an introduction to quantum mechanics and its applications to a variety of physical problems. In more advanced books, subtler and more complex problems are addressed—for example, quantum optics, the properties of many particle physics at low temperatures, and relativistic quantum mechanics of elementary particles such as quarks, gluons, and leptons. The universal experience is that in the *subclassical range*, from the *mesoscopic* (dimensions of order mm) to the *ultramicroscopic* (so far of order of 10^{-18} m), **quantum mechanics¹ works extremely well**. The relativistic quantum theory of the interaction of electrons or muons² with the electromagnetic field is incredibly successful. As an example, consider the calculated and experimental values of the values of the anomalous magnetic moments. For muons they are

$$(a_\mu)_{\text{exp}} = (116592020 \pm 150) \times 10^{-11}$$
$$(a_\mu)_{\text{theo}} = (116591708 \pm 67) \times 10^{-11}$$

In view of this and countless other examples, there is a pervasive feeling among those working in the range of mesoscopic to ultramicroscopic physics that quantum mechanics provides the correct theory of the behavior of matter and light.

Our study of quantum physics did not put much stress on the *strange* aspects of quantum mechanics. We summarize what we did point out, without great elaboration: At the start of Chapter 2 we discussed a thought experiment in which individual photons were sent through a polarizer. This led us to the introduction of the idea that the prediction of the behavior of individual photons was not possible, but that we could talk about the *probability* of what a particular photon would do. The extension of this idea to matter (electrons, neutrons, so on) led to the proposal of the existence of a wave function and its associated probability interpretation. A brief discussion of the passage of individual electrons through a pair of slits (as in the experiment of A. Tonomura discussed in Chapter 1) led us to a simple rule: If the paths of the electrons when passing through the two-slit apparatus are not determined, then there will be interference, and the probability density for finding an electron at x at the time t will be given by

$$P(x, t) = |\psi(x, t)|^2 = |\psi_1(x, t) + \psi_2(x, t)|^2 \quad (20-1)$$

¹In the subatomic domain quantum mechanics, it is necessary to use a natural extension of quantum mechanics to *relativistic quantum field theory*.

²Muons are essentially heavy electrons, and their interaction with the electromagnetic field is the same as that of electrons. Their rest mass is 207 times that of the electron rest mass, and they are unstable decaying into electrons and two (different) neutrinos, with a lifetime of the order of $2.2 \mu\text{s}$.

If, on the other hand, we find a way of determining the path of each electron, then we could, in principle, close the unused slit for each electron, and the probability density is then the sum of the individual probability densities

$$P(x, t) = |\psi_1(x, t)|^2 + |\psi_2(x, t)|^2 \quad (20-2)$$

The wave functions are solutions of the Schrödinger equation and are uniquely determined by their initial values. The equation is linear, and therefore a superposition of any number of solutions of the equation is itself a solution.

In Chapter 3 and later, we elaborated on the superposition principle by stressing the expansion postulate. We stated that any wave function $\psi(x)$ could be expanded in a complete set of eigenfunctions of any hermitian operator. Thus, if the $u_a(x)$ form a complete, orthonormal set, satisfying

$$Au_a(x) = au_a(x) \quad (20-3)$$

and

$$\int_{-\infty}^{\infty} dx u_a^*(x) u_b(x) = \delta_{ab} \quad (20-4)$$

then we could always write, for *any* square-integrable wave function,

$$\psi(x) = \sum_a C_a u_a(x) \quad (20-5)$$

The coefficients, determined using orthonormality of the eigenfunctions, were given by

$$C_a = \int_{-\infty}^{\infty} dx u_a^*(x) \psi(x) \quad (20-6)$$

We also asserted that these coefficients had the following interpretation: A measurement of the eigenvalue of A , on the system described by $\psi(x)$, must lead to one of the eigenvalues. The outcome of a particular measurement was unpredictable, with the probability of finding a particular value a given by

$$P(a) = |C_a|^2 \quad (20-7)$$

Furthermore, as a consequence of the measurement, the system is *projected* into the state described by $u_a(x)$. These statements are suitably generalized in Chapter 6 to the more abstract formulation of quantum mechanics, in which the states were described by kets, $|\psi\rangle$, $|a\rangle$, $\dots$, defined as vectors in a vector (Hilbert) space. We made extensive use of the expansion theorem in setting up perturbation theory, and there is no question as to its correctness.

Nevertheless, the assertions about the random nature of the results of the measurement, the probability interpretation of the expansion coefficients, and the projection, by a measurement, into the particular eigenstate raised questions among a number of people. These assertions *do not*, directly, follow from the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi\rangle = H|\psi\rangle \quad (20-8)$$

How to deal with these assertions is an issue concerning the foundations of quantum mechanics. I want to stress once again that, however one interprets the origin of the quantum mechanical rules, *they work*, and in the final analysis, that is all that matters.

Recent advances in technology have made it possible to carry out experiments that in the past could only be thought about. The actual execution of *thought experiments*³ confirms the validity of the rules of quantum mechanics. In one way or another the new experiments are connected with the existence of *entanglement*, the nonseparability of quantum states of systems that may not even be interacting. Before getting to entanglement, which is central to any discussion of the foundations of quantum mechanics, we first discuss the two-slit experiment again in a way that illustrates what is sometimes called the weirdness of quantum mechanics.

20-1 THE TWO-SLIT EXPERIMENT AND WHICH-WAY INFORMATION

Let us recall the two-slit experiment discussed in Chapter 2. When a beam of electrons (or photons or neutrons or, most recently, C_{60} and C_{70} molecules⁴) is incident on a two-slit system, an interference pattern, characteristic of waves, appears on the screen behind. The interference pattern observed in the passage of a particle through two slits is logically incompatible with knowledge of which slit the particle went through, since such knowledge would imply that the pattern is a superposition of the patterns created by particles coming *independently* from one slit or the other. Such a superposition of two different beams from two slits cannot give rise to an interference pattern. It is the uncertainty principle that may be used to “save” quantum mechanics from a logical self-contradiction. Imagine a monitor that is used to identify the slit of passage. Such a monitor could be a source of light, in which the scattered photon is used to determine which slit the electron comes through.

Let the slits be separated by a distance a , and the distance from the slits to the screen be denoted by D (Fig. 20-1). The condition for constructive interference is

$$\sin \theta_n = n \frac{\lambda}{a} \quad (20-9)$$

Here λ is the de Broglie wavelength of the electron, n is an integer, and θ_n is the angle between the normal to the screens and the line to the n th maximum. In this setup the distance between adjacent maxima on the screen is

$$D \sin \theta_{n+1} - D \sin \theta_n = \frac{D\lambda}{a} \quad (20-10)$$

Our monitor, in order to be effective, must determine the y -coordinate of the electron to an accuracy $\Delta y < a/2$. In doing so it imparts to the electron an *uncontrolled* amount of momentum of the order of magnitude $\Delta p_y > h/(a/2)$. This implies that

$$\frac{\Delta p_y}{p} > \frac{2h}{a(h/\lambda)} = \frac{2\lambda}{a} \quad (20-11)$$

³The German word *Gedankenexperimente* appears in much of the literature on the subject

⁴This work by A. Zeilinger and collaborators was carried out using a standing electromagnetic wave as a diffraction grating: O. Nairz et al., *Phys. Rev. Letters*, **87**, 160401 (2001).

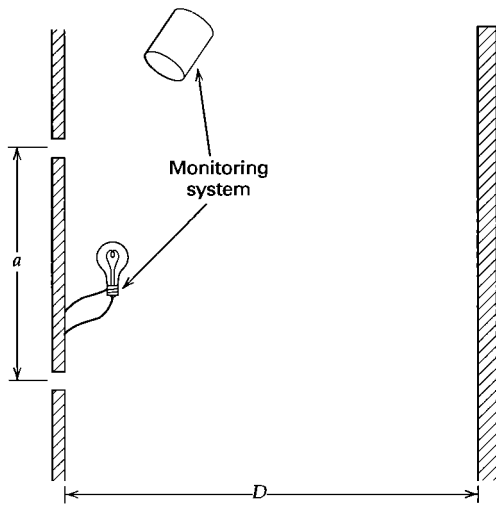


Figure 20-1 A two-slit apparatus with monitor to identify the slit of passage.

and thus the uncertainty in the location of arrival at the screen is

$$D \frac{\Delta p_y}{p} > \frac{2D\lambda}{a}$$

This displacement is larger than the distance between the adjacent maxima, and therefore wipes out the interference pattern. Quantum mechanics does not require an actual experiment to detect the slit of passage. If there is a way of, *in principle*, distinguishing the path of the electron, then the rules state that there will be no interference. A design of an experiment that allows for the possible determination of the path of the electron is called a *which-way* design. An example using photons instead of electrons follows. There is no implication that the path identification is built into the design of the experiment at the beginning. The important point is that the presence or absence of interference depends on the determination of indistinguishability or distinguishability of the paths at the end of the experiment, when the measurement occurs.

A dramatic example is a *delayed choice* experiment, first suggested by John Wheeler, and carried out with the help of a so-called Mach-Zehnder interferometer. Figure 20-2

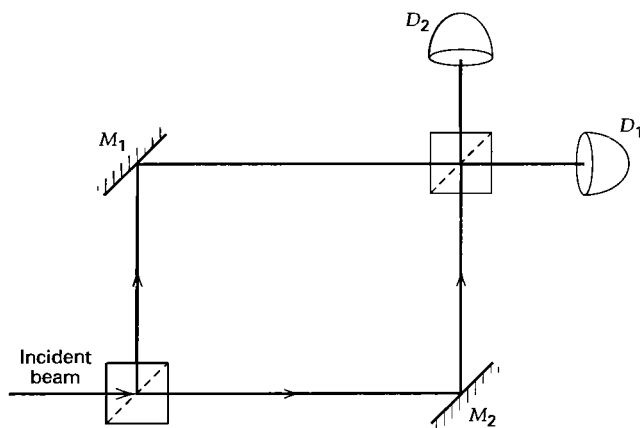


Figure 20-2 Schematic picture of Mach-Zehnder Interferometer. The boxes represent beam splitters (half-silvered mirrors), M_1 and M_2 are mirrors, D_1 and D_2 are detectors.

shows the basics of the setup, equivalent to a two-slit experiment. Individual photons enter a beam-splitter (really a half-silvered mirror) in which there is a 50% chance of the photon being deflected or transmitted. Interference is observed in both detectors. This is seen by measuring how the intensities in the two detectors vary as the mirror M_1 is moved. We may, therefore, say that each photon effectively chooses both paths. If one of the beam-splitters is removed (say, the top one), then there is no interference. A photon taking the lower path ends up in D_2 , and the photon taking the upper path ends up triggering detector D_1 . The idea of a *delayed choice* experiment is to either insert or remove the second beam-splitter after the photon has passed through the first beam-splitter. What is observed⁵ is that, if the upper beam splitter is inserted *after* the photon has passed the first beam splitter, there will be interference. If the beam-splitter is in place when the photon enters the system but is removed after the photon enters the system, then there is no interference. This is consistent with the rules we stated: If an examination of the apparatus at the end shows that there was no way of deciding which way the photon went, there will be interference.⁶ A more elaborate experiment, which also introduces the notion of a *quantum eraser*, is discussed in the next section and in Supplement 20-A [www.wiley.com/college/gasiorowicz].

20-2 THE QUANTUM ERASER

The idea of a *quantum eraser*, a mechanism by means of which *which-way* information once established can be erased, is due to M. O. Scully and K. Druhl.⁷ Supplement 20-A [www.wiley.com/college/gasiorowicz] discusses a specific mechanism designed for a two-slit experiment for atoms. Here we discuss a simpler, later version of this experiment carried out with photons.

The experiment carried out by T. J. Herzog et al.⁸ makes use of a nonlinear crystal that has the property of converting an incident photon into two lower (usually not equal) energy photons. These photons are correlated in that they are created simultaneously, and they both have the same polarization, in this case *vertical* to the plane of the experiment. The photons (named *signal* and *idler* photons) emerge at different angles to the forward direction of the incident laser beam and are reflected back, passing through the crystal and arriving at the top and bottom detectors shown in Fig. 20-3. These are the *reflected* photons, shown by solid lines in the figure. We label them by $|i_1\rangle$ and $|s_1\rangle$, respectively, upon production. They are coherent with the laser beam. The laser beam is reflected by another mirror and has a second chance to produce a two-photon pair. These photons (called the *direct* photons, since they are not reflected), labeled $|i_2\rangle$ and $|s_2\rangle$, follow the same path as the first pair upon reflection from their respective mirrors. They are indicated by the dotted lines in the figure. The two *idler* photons (frequency ω_1) are indistinguishable from each other, as are the *signal* photons (frequency ω_2). The state of the photons as they approach the detectors D_i and D_s is given by (V refers to the polarization)

$$|\Psi\rangle = C[e^{i\phi_i}|i_{1V}\rangle e^{i\phi_s}|s_{1V}\rangle + e^{i\phi_i}|i_{2V}\rangle|s_{2V}\rangle] \rightarrow C[e^{i(\phi_i+\phi_s)}|i_V\rangle|s_V\rangle] \quad (20-12)$$

⁵This has been established by P. Grangier, G. Roger, and A. Aspect in 1986. The beam-splitter is not really inserted physically, since the time interval for doing so is too short. A device that can respond in such a short time effectively simulates the insertion and removal of the beam-splitter.

⁶A more detailed discussion of this and many other experiments probing the fundamentals of quantum mechanics may be found in the very interesting book *The Quantum Challenge*, by G. Greenstein and A. G. Zajonc, Jones and Bartlett Publishers, Sudbury, Mass., 1997.

⁷M. O. Scully and K. Druhl, *Phys. Rev. A*, **25**, 2208 (1982).

⁸T. J. Herzog, P. G. Kwiat, H. Weinfurter, and A. Zeilinger, *Phys. Rev. Letters*, **75**, 3034 (1995).

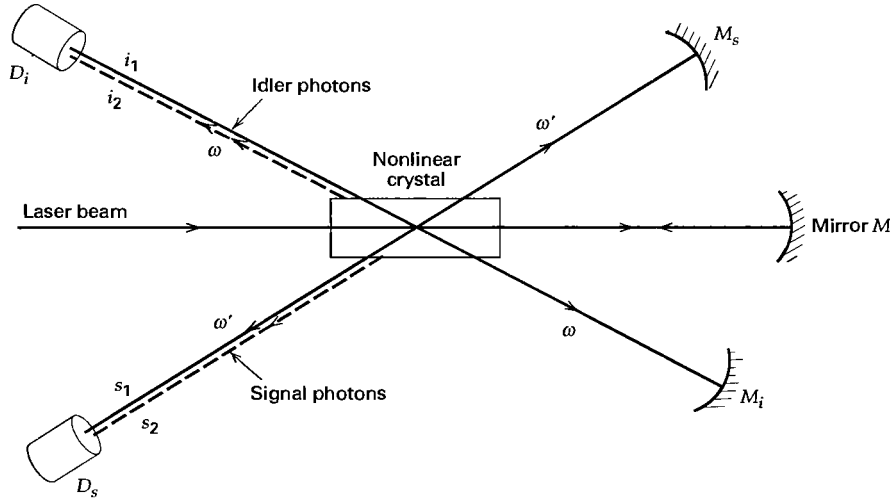


Figure 20-3 Schematic picture of a quantum eraser based on set up by T. J. Herzog, et al., *Phys. Rev. Letters*, **75**, 3034 (1995).

The phase factors have to do with the fact that the photons “1” were reflected by the mirrors by the time they reached the detector. The photons “2” were created by the reflected laser beam, which also undergoes a phase change. Since the photons i_1 and i_2 are identical, we label them by $|i\rangle$ and similarly for the signal photons. We will continue to drop the “1” and “2” labels when the photons (idler/signal) are indistinguishable. The intensities at the two detectors are identical, and are proportional to

$$|e^{i(\phi_i + \phi_s)} + e^{i\phi_L}|^2 = 2(1 + \cos(\phi_i + \phi_s - \phi_L)) \equiv 2(1 + \cos \Delta\phi) \quad (20-13)$$

The oscillatory pattern of intensity has been observed by moving the detectors or mirrors along the beam axis. (Note that the i and s mirrors have to be moved in opposite directions to maintain the same intensity, since these phases appear as a sum in $\Delta\phi$). The interference mechanism is equivalent to a two-slit experiment. The which-way effect can be seen when the upper mirror M_s is removed. In that case there is no signal (reflected) photon associated with the reflected idler photon, which allows us to distinguish between the two idler photons arriving at the detector D_i . A coincidence measurement would show that when D_s clicks, then the click in D_i had to be an i_2 photon. This is analogous to knowing which of the two slits a photon passes through. The interference disappears. Next, the mirror is put back, and a plate that rotates the polarization from V to H is placed in front of the lower mirror M_i . As expected, the interference pattern disappears at the detector D_i , since the two idler photons now have orthogonal polarizations. The action of the plate is shown here:

$$|\Psi\rangle = C[e^{i\phi_i}|i_{1V}\rangle e^{i\phi_s}|s_{1V}\rangle + e^{i\phi_L}|i_{2V}\rangle|s_{2V}\rangle] \rightarrow C[e^{i(\phi_i + \phi_s)}|i_{1H}\rangle|s_{1V}\rangle + e^{i\phi_L}|i_{2V}\rangle|s_{2V}\rangle] = C[e^{i(\phi_i + \phi_s)}|i_{1H}\rangle + e^{i\phi_L}|i_{2V}\rangle]|s_V\rangle \quad (20-14)$$

The absolute square of this shows no interference terms, since $\langle i_V|i_H\rangle = 0$. Since this allows us *in principle* to tell the two photons apart (it is possible to set up a polarization experiment after reflection and before the reflected idler photon passes through the crystal, but one does not actually have to do so), we have a which-way situation. Distinguishing the reflected idler photon then distinguishes the reflected signal photon, since these are always produced in pairs. Thus there will be no interference in the signal detector D_s !

If we express the photon states not in terms of H and V but in terms of the states rotated through $\pm 45^\circ$,

$$|H\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle); |V\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle) \quad (20-15)$$

then the previous state may be written as

$$e^{i(\phi_i + \phi_s)}|i_H\rangle|s_V\rangle + e^{i\phi_i}|i_V\rangle|s_V\rangle = \frac{1}{\sqrt{2}}\{e^{i(\phi_i + \phi_s)}(|i_+\rangle + |i_-\rangle)|s_V\rangle + e^{i\phi_i}(|i_+\rangle - |i_-\rangle)|s_V\rangle\} \quad (20-16)$$

The next step involves putting a 45° filter in front of the detector D_i , while keeping the plate in front of the mirror M_i . As a consequence the state $|i_-\rangle$ is absorbed. Effectively, the idler photons are described by $(|V\rangle + |H\rangle)/\sqrt{2}$, and it is no longer possible to tell whether a registered idler photon was originally V or H . This corresponds to a *quantum eraser* in that former which-way information has been erased. The state becomes

$$\frac{1}{\sqrt{2}}\{e^{i(\phi_i + \phi_s)}|i_+\rangle|s_V\rangle + e^{i\phi_i}(|i_+\rangle)|s_V\rangle\} \rightarrow \frac{1}{\sqrt{2}}(e^{i\Delta\phi} + 1)|i_+\rangle|s_V\rangle \quad (20-17)$$

The same interference pattern as in (20-13) appears.

A final experiment takes the preceding setup and, to begin with, places a polarization-rotating plate in front of the mirror M_s . This rotates the reflected signal photon's polarization from V to H . Since the direct and reflected photons are now distinguished, we have to write the effect of this on the state in (20-17) as follows:

$$\frac{1}{\sqrt{2}}\{e^{i(\phi_i + \phi_s)}|i_+\rangle|s_{1V}\rangle + e^{i\phi_i}(|i_+\rangle)|s_{2V}\rangle\} \rightarrow \frac{1}{\sqrt{2}}|i_+\rangle(e^{i\Delta\phi}|s_{1H}\rangle + |s_{2V}\rangle) \quad (20-18)$$

In this form, because of the polarization distinction between the two signal photons, there will be no interference. We rewrite the above in the form

$$\begin{aligned} \frac{1}{\sqrt{2}}|i_+\rangle(e^{i\Delta\phi}|s_{1H}\rangle + |s_{2V}\rangle) &= \frac{1}{2}|i_+\rangle\{e^{i\Delta\phi}(|s_{1+}\rangle + |s_{1-}\rangle) + (|s_{2+}\rangle - |s_{2-}\rangle)\} \\ &= \frac{1}{2}|i_+\rangle\{|s_+\rangle(e^{i\Delta\phi} + 1) + |s_-\rangle(e^{i\Delta\phi} - 1)\} \end{aligned} \quad (20-19)$$

The next step is to put a polarization rotator in front of the signal detector. If we absorb the $|s_-\rangle$ photon, then the interference pattern is the same as in (20-13). If we absorb the $|s_+\rangle$ photon, we get the complementary interference pattern

$$|e^{i\Delta\phi} - 1|^2 = 2(1 - \cos \Delta\phi) \quad (20-20)$$

We see that the polarization rotator acts as a quantum eraser. The absence of polarization can be seen as a sum of two complementary interference patterns.⁹ This is apparent in the data of the experiment of Herzog et al.

20-3 ENTANGLEMENT AND EPR

In our discussion of quantum mechanics we paid very little attention to a particular property of quantum mechanical states, which we may roughly describe as nonseparability. In our discussion of two-spin states in Chapter 10, we did note that in a singlet state of the form

$$X_{\text{singlet}} = \frac{1}{\sqrt{2}}(\chi_{\uparrow}^{(1)}\chi_{\downarrow}^{(2)} - \chi_{\downarrow}^{(1)}\chi_{\uparrow}^{(2)}) \quad (20-21)$$

⁹A similar pattern is seen in the experiment discussed in Supplement 20-A [www.wiley.com/college/gasiorowicz].

if particle (1) was found to be in an eigenstate of $S_z^{(1)}$ with eigenvalue $\pm\hbar/2$, the particle (2) would be found in an eigenstate of $S_z^{(2)}$ with eigenvalue *necessarily* equal to $\mp\hbar/2$. We also found that if a measurement of the x -component of the spin of particle (1) gave an eigenvalue of $\pm\hbar/2$ (each with probability $1/2$), then a measurement of the x -component of the spin of particle (2) would always give the opposite sign eigenvalue. This nonseparability does not apply to a simple product of states. For a state described by $\chi_{\uparrow}^{(1)} \chi_{\downarrow}^{(2)}$ measurements of the x -components would give opposite sign eigenvalues, but even if a measurement of the x -component of the spin of particle (1) gave a positive eigenvalue, a measurement of the x -component of the spin of particle (2) would have an equal chance of being positive or negative. The state described in (20-21) is said to be *entangled*. The measurement carried out on particle (1) does not result in a communication from that particle to particle (2). The two measurements of the x -component of the spins can be carried out *simultaneously*—that is, when the separation between the locations of the particles is *spacelike*. Entanglement plays a very special role in quantum physics, and its existence leads to a number of counterintuitive conclusions. In 1935 in a paper by Einstein, Podolsky and Rosen (EPR) a seemingly plausible requirement for a physical theory was set forth. The criteria, of which we list just two, contain the following.¹⁰

1. Every element of physical reality must have a counterpart in a *complete* physical theory.
2. If without disturbing the system we can predict, with certainty, the value of some observable, then we can associate an element of reality with this observable.

EPR analyzed a hypothetical experiment in their paper, but their argument that quantum theory is not a *complete* theory can easily be expressed in terms of the two-spin system, as was suggested by Bohm. Suppose we measure the z -component of the spin of particle (1) and we find that the eigenvalue is $+\hbar/2$. We can then predict, without in any way disturbing the system, that the z -component of the spin of particle (2) will be $-\hbar/2$. According to EPR, we may thus associate with the z -component of spin of particle (2) an element of physical reality, meaning that particle (2) must have had $S_z = -\hbar/2$ all along, since the measurement on particle (1)—undertaken at an arbitrary distance from where particle (2) is being measured—could not possibly affect its state. Suppose now that, instead of measuring the z -component of the spin of particle (1), the x -component were measured. The same argument would go through, and thus one could ascribe to particle (2) both a value of S_z and S_x . This is incompatible with quantum theory. According to the EPR criteria, either quantum mechanics is not a *complete* theory or there must exist a *nonlocal* interaction between the two particles. The nonlocality does not contradict special relativity, since no message can be sent from one particle to the other when these *incompatible* experiments are performed. Bohr responded to this criticism by asserting that quantum mechanics was different from classical physics, in that it was meaningless to try to simultaneously assign values to physical quantities represented by operators that do not commute. In this example S_z and S_x do not commute. This was an example of what he called *complementarity*,¹¹ a view that quantum mechanical systems may have aspects that cannot all be determined simultaneously (as, for example, particle/wave properties). Nevertheless, a number of people were not satisfied by this explanation. A

¹⁰Here we follow the discussion given in the fine book *Quantum Theory* by David Bohm, Dover Publications, New York, 1989.

¹¹This is a very restricted application of the general approach of Bohr to the deep differences between classical and quantum physics.

way out would be the existence of certain *hidden variables*, which would in some sense describe a linkage between the two *separate* states. This modification of the theory would have to be such as not to affect the observed successes of quantum mechanics. An important step forward in the discussion of this program was due to John Bell. Bell¹² analyzed the requirements for a *complete* and *local* theory, and studied correlations between spin measurements of the separate particles in different directions. These were found to be subject to certain inequalities that contradicted what would be expected from quantum mechanics. Subsequent experiments¹³ showed complete agreement with the expectations of quantum mechanics, which rely purely on the existence of *entangled states* of the form described in (20-21). The result of these beautiful and ingenious experiments should not have come as a surprise: A variety of correlation experiments in nuclear and atomic physics never contradicted quantum mechanical expectations, and although the particular correlation suggested by Bell had not been measured before, there was no reason to expect trouble.

A more dramatic confrontation between EPR realism and quantum mechanics was worked out by Greenberger, Horne, and Zeilinger (GHZ).¹⁴ We have three measuring recorders, and each one has two settings, which are labeled *A* and *B*. The three recorders are all spacelike with respect to each other, and they record events by an “up” or “down” result. There will therefore be records that have the label $A_1, A_2, A_3, B_1, B_2, B_3$, and each will give a result $+1$ or -1 . On the average, any one of the readings A_i half the time will be positive and half the time negative, and similarly for the B_i readings. The experimenters however find certain correlations, in that certain products are not random. They find that whatever the readings are, the following is true:

$$A_1 B_2 B_3 = B_1 A_2 B_3 = B_1 B_2 A_3 = +1 \quad (20-22)$$

Since these recorders are spacelike with respect to each other, they cannot affect each other, and therefore correlations must come about as a result of some hidden variables. One consequence of the above readings, combined with the fact that

$$B_1^2 = B_2^2 = B_3^2 = 1 \quad (20-23)$$

is that

$$A_1 A_2 A_3 = +1 \quad (20-24)$$

In fact, the results described in (20-22) are just the consequences of measurements of the observables

$$A_i = \sigma_x^{(i)}; \quad B_i = \sigma_y^{(i)} \quad (20-25)$$

Here $\sigma_{x,y}^{(i)} = 2/\hbar S_{x,y}^{(i)}$ are operators with eigenvalues ± 1 , that act on the entangled three-particle state

$$\Psi(1, 2, 3) = \frac{1}{\sqrt{2}} (\chi_{\uparrow}^{(1)} \chi_{\uparrow}^{(2)} \chi_{\uparrow}^{(3)} - \chi_{\downarrow}^{(1)} \chi_{\downarrow}^{(2)} \chi_{\downarrow}^{(3)}) \quad (20-26)$$

¹²Bell's theorem is treated in many textbooks (see the books by Peebles, Park, Sakurai, Griffiths, and Ohanian, listed in the bibliography). See also *The Quantum Challenge*, loc. cit.

¹³The experiments by A. Aspect and collaborators are discussed in *The Quantum Challenge*, loc. cit.

¹⁴For a detailed description see D. Greenberger, M. Horne, A. Shimony, and A. Zeilinger, *Am. J. Phys.* **58** (1990) 1131.

It follows from

$$\begin{aligned}
 \sigma_x \chi_\uparrow &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \chi_\downarrow \\
 \sigma_x \chi_\downarrow &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \chi_\uparrow \\
 \sigma_y \chi_\uparrow &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix} = i\chi_\downarrow \\
 \sigma_y \chi_\downarrow &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ 0 \end{pmatrix} = -i\chi_\uparrow
 \end{aligned} \tag{20-27}$$

that

$$\begin{aligned}
 A_1 \Psi(1, 2, 3) &= \frac{1}{\sqrt{2}} (\chi_\downarrow^{(1)} \chi_\uparrow^{(2)} \chi_\uparrow^{(3)} - \chi_\uparrow^{(1)} \chi_\downarrow^{(2)} \chi_\uparrow^{(3)}) \\
 A_2 \Psi(1, 2, 3) &= \frac{1}{\sqrt{2}} (\chi_\uparrow^{(1)} \chi_\downarrow^{(2)} \chi_\uparrow^{(3)} - \chi_\downarrow^{(1)} \chi_\uparrow^{(2)} \chi_\uparrow^{(3)}) \\
 A_3 \Psi(1, 2, 3) &= \frac{1}{\sqrt{2}} (\chi_\uparrow^{(1)} \chi_\uparrow^{(2)} \chi_\downarrow^{(3)} - \chi_\downarrow^{(1)} \chi_\downarrow^{(2)} \chi_\uparrow^{(3)}) \\
 B_1 \Psi(1, 2, 3) &= \frac{i}{\sqrt{2}} (\chi_\downarrow^{(1)} \chi_\uparrow^{(2)} \chi_\uparrow^{(3)} + \chi_\uparrow^{(1)} \chi_\downarrow^{(2)} \chi_\uparrow^{(3)}) \\
 B_2 \Psi(1, 2, 3) &= \frac{i}{\sqrt{2}} (\chi_\uparrow^{(1)} \chi_\downarrow^{(2)} \chi_\uparrow^{(3)} + \chi_\downarrow^{(1)} \chi_\uparrow^{(2)} \chi_\uparrow^{(3)}) \\
 B_3 \Psi(1, 2, 3) &= \frac{i}{\sqrt{2}} (\chi_\uparrow^{(1)} \chi_\uparrow^{(2)} \chi_\downarrow^{(3)} + \chi_\downarrow^{(1)} \chi_\downarrow^{(2)} \chi_\uparrow^{(3)})
 \end{aligned} \tag{20-28}$$

Consequently, for example,

$$B_1 B_2 A_3 \Psi(1, 2, 3) = \frac{i^2}{\sqrt{2}} (\chi_\downarrow^{(1)} \chi_\uparrow^{(2)} \chi_\downarrow^{(3)} - \chi_\uparrow^{(1)} \chi_\uparrow^{(2)} \chi_\uparrow^{(3)}) = \Psi(1, 2, 3) \tag{20-29}$$

and so on. However, in the application of $A_1 A_2 A_3$ each of the spins is flipped (without the i or $-i$ coefficients), so that the quantum mechanical prediction is that

$$A_1 A_2 A_3 \Psi(1, 2, 3) = \frac{1}{\sqrt{2}} (\chi_\downarrow^{(1)} \chi_\downarrow^{(2)} \chi_\downarrow^{(3)} - \chi_\uparrow^{(1)} \chi_\uparrow^{(2)} \chi_\uparrow^{(3)}) = -\Psi(1, 2, 3) \tag{20-30}$$

in complete contradiction with the local theory that does not include the consequences of entanglement. Recent advances in quantum optics have made it possible to actually construct GHZ states for photons and check the above result. (See Supplement 20-B [www.wiley.com/college/gasiorowicz].)

20-4 QUANTUM MEASUREMENTS¹⁵

In most areas of physics there is very little discussion of *measurement* as part of the theory. It is different in quantum mechanics. As was illustrated in our examples of which-way and delayed-choice experiments, the process of measurement and the design of the

¹⁵In this brief discussion, I lean for much of the material on N. G. Van Kampen, *Physica A* **153**, 97 (1988), and on discussions in Roland Omnès, *Understanding Quantum Mechanics*, Princeton University Press, Princeton, N.J., 1999. The books by D. Bohm, P. J. Peebles, and especially K. Gottfried deal with this subject. Since a deeper discussion of this material involves the *density matrix*, we provide a brief description of the density matrix formalism in Supplement 20-C [www.wiley.com/college/gasiorowicz].

apparatus determine whether a photon acts as a particle or as a wave, for example. The problem of quantum measurements has attracted a certain amount of attention in recent years, and although the people working in this field have not agreed on the matter, the problem deserves a brief discussion.

Matter is made of atoms, and atoms obey quantum mechanics. Thus, in principle, a description of a measurement of a microscopic system requires that the macroscopic detector also be described by a state vector. As we shall see, this necessarily brings in the notion of entanglement of the system and the apparatus. We will give a qualitative discussion of a quantum measurement.

For our purposes, it is simplest to consider a measurement of the spin of a neutral atom of spin $1/2$. The initial spin-state may be described by

$$|\psi\rangle = \alpha|+\rangle + \beta|-\rangle \quad (20-31)$$

The apparatus is described by the state vector $|\Phi(0)\rangle$, so that initially the total state vector for the system is

$$|\Psi(0)\rangle = |\psi\rangle|\Phi(0)\rangle \quad (20-32)$$

The state of the apparatus is very complicated. Although we can talk about a *pointer* that can read differently for spin-up and spin-down, the pointer is described by many degrees of freedom. The atom enters the region of the apparatus, and the consequent development of the system is described by quantum mechanics. In particular, we have

$$|\Psi(t)\rangle = e^{-iHt/\hbar}|\Psi(0)\rangle \quad (20-33)$$

Here H is the total Hamiltonian for the system. It describes the motion of the atom, and it describes the interaction between the atom and the apparatus. We will set up the apparatus in such a way that it does not alter the initial state $|\psi\rangle$. For example, a Stern-Gerlach experiment does not rotate the spin; it merely changes the direction in which the atoms move, with the direction dependent on the spin state. The consequence of the interaction, as described by (20-33), is that

$$|\Psi(0)\rangle \rightarrow |\Psi(t)\rangle = \alpha|+\rangle|\Phi_{\text{up}}\rangle + \beta|-\rangle|\Phi_{\text{down}}\rangle \quad (20-34)$$

The apparatus and the microscopic spin system have become entangled. Note that this time evolution does *not* describe a measurement. We still have a superposition of two state vectors. In principle, it would be possible by an arrangement of magnets to reverse the effect of the first set of magnets and get back to the initial state. We need to do something that actually distinguishes the $|\Phi_{\text{up/down}}\rangle$ states, which differ in the physical location of the center of mass of the atoms. A position measurement of the atoms could be accomplished by a two-slit addition to the apparatus, with a monitor to determine which slit the atom went through. The state vector, which evolves according to (20-33), now includes the state of the slit-detector, and it becomes

$$|\Psi(t)\rangle = \alpha|+\rangle|\Phi_{\text{up}}\rangle|S_{\text{upper}}\rangle + \beta|-\rangle|\Phi_{\text{down}}\rangle|S_{\text{lower}}\rangle \quad (20-35)$$

This, however, is still a coherent superposition of two states. What about the step of observing the monitor? We could go on and on with our quantum mechanical evolution, but finally something *irreversible* must happen. In our example, the action of the monitor that records which slit the atom went through must leave a permanent record somewhere. For example, the monitor gives rise to a current, and the current readings are recorded on a disk in a computer.

There are several questions that come up: (1) Does one ever see linear combinations of *macroscopic* systems such as shown in (20-34) or (20-35), and (2) if not, why not?

(3) How does one end up with the system *projected* into the state indicated by the measurement; that is to say, how does the transition to $\alpha |+\rangle |\Phi_{\text{up}}\rangle |S_{\text{upper}}\rangle$ or $\beta |-\rangle |\Phi_{\text{down}}\rangle |S_{\text{lower}}\rangle$ come about? With this transition, since the apparatus state vectors are separately normalized, we find that the spin-up and spin-down states of the atom occur with probabilities $|\alpha|^2$ and $|\beta|^2$, respectively.

Decoherence and the Schrödinger Cat

The idea of a linear superposition of macroscopic states seems preposterous. Schrödinger made this very vivid by proposing an experiment in which the macroscopic *pointer* is a cat. For example, consider the two slits as being small holes in a closed box that holds a live cat. If the atom goes through the upper hold, it triggers a device that opens a vial of cyanide and kills the cat. If the atom goes through the lower hole, nothing happens. Until a final measurement is made, we are presented with a situation in which we have an entangled state of the atom path (upper and lower) and the state of the cat. This state can be written in the form

$$|\Psi\rangle = \alpha |+\rangle |\text{dead}\rangle + \beta |-\rangle |\text{live}\rangle \quad (20-36)$$

At some point, when the box is opened, one finds that the cat is either alive or dead. When did the “projection” of the combined state into the definite state occur? Was a measurement—that is, the opening of the box—really necessary?

The commonsense answer is that as soon as the atom, with probability $|\alpha|^2$, triggered the breaking of the vial, the cat died, and that brings back question (2) above. Still, would a superposition of the form (20-36) ever occur? Until recently the answer would have been that this does not make sense. However, recent experiments have shown that relatively large systems, *mesoscopic* systems, do appear in linear combinations. The simplest example we can quote is the two-slit experiment with fullerenes (C_{70}), mentioned in footnote 4. The authors propose doing a similar experiment with a virus! A more dramatic experiment involves superconducting quantum interference devices (SQUID). Basically one deals with superconducting states that are macroscopic quantum systems. Under certain circumstances, pairs of electrons can become correlated. These pairs form bosons, and at very low temperatures bosons preferentially all go into the same state, which becomes what is called a *condensate*, a single quantum system. Such a system may involve up to 10^{10} pairs, and comes close to what would be called a macroscopic system. Linear combinations of two such different condensates have been observed.¹⁶ In effect, the technology is there to show that linear superpositions are not limited to electrons and atoms.

There is no chance of ever doing the cat experiment. The reason is a phenomenon that has received a lot of attention lately, called *decoherence*.¹⁷ The point is that all of our discussion has dealt with *closed* systems, in which an equation like (20-33) describes their evolution. Actually, unless we are talking about the whole universe, any system that we study is, to some extent, an *open* system that interacts with an *environment*. The implications for this are that a state like (20-34), for example, ought to be written in the form

$$|\Psi(0)\rangle \rightarrow |\Psi(t)\rangle = \alpha |+\rangle |\Phi_{\text{up}}, \eta_{\text{up}}\rangle + \beta |-\rangle |\Phi_{\text{down}}, \eta_{\text{down}}\rangle \quad (20-37)$$

¹⁶For a description, see a brief article by G. Blatter in *Nature*, **406**, 25 (2000). The journal *Physics World* has frequent articles on such interesting experiments.

¹⁷There is extensive literature on this subject. A readable introduction to the subject appears in W. H. Zurek, *Physics Today*, **44**, 10, p. 36 (1991).

The variables denoted by η_{up} , η_{down} describe symbolically different environments. If we now ask for the expectation value of an operator that mixes the *up* and *down* states—for example, σ_x —then

$$\begin{aligned}\langle\sigma_x\rangle &= \{\alpha^*\langle +|\langle\Phi_{\text{up}}, \eta_{\text{up}}| + \beta^*\langle -|\langle\Phi_{\text{down}}, \eta_{\text{down}}|\}\sigma_x\{\alpha|+\rangle|\Phi_{\text{up}}, \eta_{\text{up}}\rangle + \beta|-\rangle|\Phi_{\text{down}}, \eta_{\text{down}}\rangle\} \\ &= \alpha^*\beta\langle\Phi_{\text{up}}, \eta_{\text{up}}|\Phi_{\text{down}}, \eta_{\text{down}}\rangle + c. c.\end{aligned}\quad (20-38)$$

while

$$\begin{aligned}\langle\sigma_z\rangle &= \{\alpha^*\langle +|\langle\Phi_{\text{up}}, \eta_{\text{up}}| + \beta^*\langle -|\langle\Phi_{\text{down}}, \eta_{\text{down}}|\}\sigma_z\{\alpha|+\rangle|\Phi_{\text{up}}, \eta_{\text{up}}\rangle + \beta|-\rangle|\Phi_{\text{down}}, \eta_{\text{down}}\rangle\} \\ &= |\alpha|^2\langle\Phi_{\text{up}}, \eta_{\text{up}}|\Phi_{\text{up}}, \eta_{\text{up}}\rangle - |\beta|^2\langle\Phi_{\text{down}}, \eta_{\text{down}}|\Phi_{\text{down}}, \eta_{\text{down}}\rangle \\ &= |\alpha|^2 - |\beta|^2\end{aligned}\quad (20-39)$$

Because there are many *up* and *down* environment variables, and because the phases between them are uncorrelated, then averaging out the environment variables leads to the vanishing of the expectation value in (20-38). What this implies is that we get the same result as if we had the situation of two sets of states, $|+\rangle$ and $|-\rangle$, not in any linear combination but coming out of separate “boxes” with relative frequencies $|\alpha|^2$ and $|\beta|^2$. In effect, decoherence makes things *look as if* we were ending up with the projected states.

Note, however, that we have not devised a mechanism to actually carry out the projection. The time evolution in (20-37) still maintains the linear combination of states, and there is no way of getting around that. In effect, there is no completely satisfactory answer to question (2). This has stimulated a variety of paths toward either a change in quantum mechanics, or a better interpretation.¹⁸ My own view is that until now, none of the proposed changes in quantum mechanics have reproduced all of its successes¹⁹ and led to anything new, other than a more satisfying formulation of the subject from the point of view of the authors.²⁰ Much of the work in this field has become rather technical, and so it is not appropriate to give a critical summary at the level of this book.²¹

¹⁸A critique of the “Copenhagen Interpretation” appears in a number of papers by J. S. Bell, collected in J. S. Bell, *Speakable and Unsayable in Quantum Mechanics*, Cambridge University Press, Cambridge, England, 1987. See also R. Omnes, *loc. cit.*

¹⁹Unless they have been designed to always give the same results as the Schrödinger equation!

²⁰The most fully developed *hidden variable* theory designed to reproduce all the successes of nonrelativistic quantum mechanics, while dealing with the interpretation issues is that due to the late David Bohm. The program is laid out in D. Bohm and B. J. Hiley *The Undivided Universe*, Routledge, London (1993).

²¹The reader will have noticed that, in contrast to most of the chapters, there are many references to books and papers. The purpose is to allow the interested reader to follow up on the technical discussions of various topics. The list is limited and selective, but many papers can be traced by starting from the references given here.

Supplement 20-A

Which-way Measurements and the Quantum Eraser¹

Consider a beam of atoms passing through a double-slit arrangement. In the absence of any attempt to gain which-way information, the atoms will create an interference pattern on the screen. The wave function is

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2}} (\psi_1(\mathbf{r}) + \psi_2(\mathbf{r})) \quad (20A-1)$$

The two parts of the wave function are those that would appear if slits 2 and 1, respectively, were closed. The probability density for finding an atom at a point $\mathbf{R}$ on the screen is

$$P(\mathbf{R}) = \frac{1}{2} |\psi_1(\mathbf{R}) + \psi_2(\mathbf{R})|^2 = \frac{1}{2} (|\psi_1(\mathbf{R})|^2 + |\psi_2(\mathbf{R})|^2 + \psi_1^*(\mathbf{R})\psi_2(\mathbf{R}) + \psi_2^*(\mathbf{R})\psi_1(\mathbf{R})) \quad (20A-2)$$

which shows the interference term. Let us next consider a way of implementing a which-way detection scheme. The proposal for path detection is quite subtle. Since we are dealing with a beam of atoms, we may excite them in a well-defined way by a carefully chosen laser beam, which crosses their path before they enter the region of the slits. The only difference is that the atomic wave function, in which $\mathbf{r}$ describes the center of mass of the atom, now has a label on it, so that

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2}} (\psi_1^{(a)}(\mathbf{r}) + \psi_2^{(a)}(\mathbf{r})) \quad (20A-3)$$

The label a identifies the electronic state of the atom. In the proposed experiment, microwave cavities are placed in front of the two slits (Fig. 20A-1). Atoms that pass through one or the other of the cavities will make a transition to a lower state of excitation. The authors deal with atoms of rubidium, with possible transitions from $(n = 63)p_{3/2}$ to $(n = 61)d_{5/2}$ or $(n = 61)d_{3/2}$. Such a transition, accompanied by the spontaneous emission of a photon, will be labeled by $a \rightarrow b$ in our formulas. This means that, depending on the path of the atom, one or the other of the cavities will now contain a photon. The wave function now becomes

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2}} (\psi_1^{(b)}\chi_1^{(1)}\chi_2^{(0)} + \psi_2^{(b)}\chi_1^{(0)}\chi_2^{(1)}) \quad (20A-4)$$

¹M. O. Scully, B.-G. Englert, and H. Walter, *Nature*, **351**, 111 (1991). This is discussed in detail in *Quantum Optics* by M. O. Scully and M. S. Zubairy, Cambridge University Press, Cambridge, England, 1997, in Chapter 20.

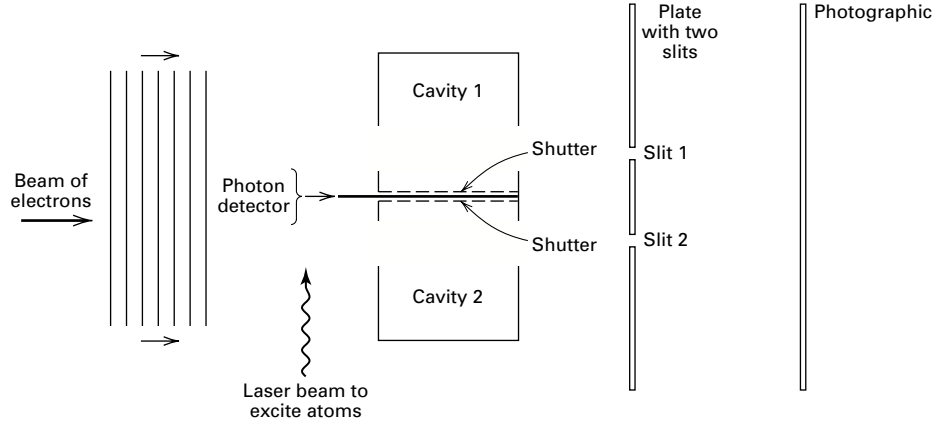


Figure 20A-1 Schematic picture of quantum eraser for atoms as described by M. O. Scully, et al. *Nature*, **351**, 111 (1991).

The lower label on $\chi_i^{(n)}$ ($i = 1, 2$) labels the cavity, and the upper one labels the number of photons in that cavity. If we now look at $|\Psi(\mathbf{R})|^2$, we see that as a consequence of the orthogonality of cavity states with zero or one photon,

$$\chi_i^{(m)*} \chi_i^{(n)} = \delta_{mn} \quad (i = 1, 2) \quad (20A-5)$$

the interference terms disappear. Note that the disappearance of the interference terms arises because we can distinguish between the cavity states, and these are *entangled* with the states of the atom. There is no uncontrollable momentum transfer. If the cavities are not empty but contain many photons, then the appearance of one more photon is not distinguishable, and under those circumstances there is no which-way detection, and the interference remains.

The fascinating aspect of the paper is the notion that the information obtained by the photon presence in one or other of the cavities can be *erased* at some later time and the interference reappears. Consider the apparatus modified in such a way that a detector is placed between the two cavities, with shutters separating the cavities and the detector. When the shutters both open, the photon in the cavity is absorbed by the detector, and then all knowledge of the photon's location is erased. One expects that the interference pattern can be re-established. Since the opening of the shutters can take place long after the photons hit the screen, we need to answer the question: How does one regain the interference pattern? The wave function now has an additional component that describes the state of the detector. When the shutters are opened, the detector changes its state from the ground state ϕ_0 to the excited state ϕ_e . The shutters are so arranged that we cannot tell whether the photon came from cavity "1" or cavity "2." This symmetry is important, since otherwise we would not lose the which-way information. To make explicit use of this symmetry, we write the wave function $\Psi(\mathbf{r})\phi_0$ by making use of the symmetric and anti-symmetric combinations

$$\psi_{\pm}^{(b)}(\mathbf{r}) = \frac{1}{\sqrt{2}} (\psi_1^{(b)}(\mathbf{r}) \pm \psi_2^{(b)}(\mathbf{r}))$$

and

$$\xi_{\pm} = \frac{1}{\sqrt{2}} (\chi_1^{(1)} \chi_2^{(0)} \pm \chi_1^{(0)} \chi_2^{(1)})$$

In terms of these the wave function reads

$$\Psi(\mathbf{r})\phi_0 = \frac{1}{\sqrt{2}} (\psi_+^{(b)}\xi_+\phi_0 + \psi_-^{(b)}\xi_-\phi_0) \quad (20A-6)$$

When the shutters are opened, the photon is absorbed. Because of the symmetry under the interchange “1” $\leftrightarrow$ “2,” the two terms behave differently when the photon is absorbed. In the first term $\chi_1^{(i)}\phi_0 \rightarrow \chi_0^{(i)}\phi_e$, and as a consequence $\chi_+\phi_0 \rightarrow \chi_0^{(1)}\chi_0^{(2)}\phi_e$. The term involving χ_- does not change, since it is antisymmetric under the interchange “1” $\leftrightarrow$ “2.” This means that after the opening of the shutters the wave function becomes

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2}} (\psi_+^{(b)}(\mathbf{r})\chi_0^{(1)}\chi_0^{(2)}\phi_e + \psi_-^{(b)}(\mathbf{r})\chi_-\phi_0) \quad (20A-7)$$

Let us now ask for the probability density at the screen, when $\mathbf{r} = \mathbf{R}$. Taking the absolute square of (20A-7), we get

$$|\Psi(\mathbf{R})|^2 = \frac{1}{2} (|\psi_+(\mathbf{R})|^2 + |\psi_-(\mathbf{R})|^2) = \frac{1}{2} (|\psi_1^{(b)}(\mathbf{R})|^2 + |\psi_2^{(b)}(\mathbf{R})|^2) \quad (20A-8)$$

There are no interference terms. The point is that we have not checked whether the which-way information has really disappeared. To do this, we must look at the detector and correlate that information with the atoms hitting the screen. The authors propose that we look at the atoms as they hit the screen one by one, and in each case ask whether the detector was in the excited state or the ground state. If it was in the excited state, then we square the part of the wave function that multiplies ϕ_e , and we get

$$P_e(\mathbf{R}) = \frac{1}{2} |\psi_+(\mathbf{R})|^2 = \frac{1}{4} (|\psi_1^{(b)}(\mathbf{R})|^2 + |\psi_2^{(b)}(\mathbf{R})|^2 + 2 \operatorname{Re} \psi_1^{(b)*}(\mathbf{R})\psi_2^{(b)}(\mathbf{R})) \quad (20A-8)$$

In the same way, the probability density for finding the atom at the screen *while the detector is in its ground state* is

$$P_0(\mathbf{R}) = \frac{1}{2} |\psi_-(\mathbf{R})|^2 = \frac{1}{4} (|\psi_1^{(b)}(\mathbf{R})|^2 + |\psi_2^{(b)}(\mathbf{R})|^2 - 2 \operatorname{Re} \psi_1^{(b)*}(\mathbf{R})\psi_2^{(b)}(\mathbf{R})) \quad (20A-9)$$

Figure 20A-2 gives a plot of the two terms. How can we say that with the *quantum eraser* in position, the fringes reappear? Let us follow the course of an atom through the apparatus, and note that it appears on the screen. We now open the shutters and see whether the detector actually absorbs a photon. If that is the case we know that the evidence of a photon has been erased. We then call this a *red* atom, and we know that it should belong to the distribution $P_e(\mathbf{R})$. For the *red* atoms, the which-way information has been lost. After we follow another atom we may find that the detector is in its ground state, so that no photon has been absorbed. This atom would belong to the class

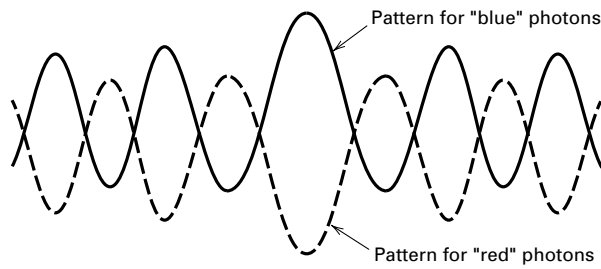


Figure 20A-2 Reappearance of fringes with quantum eraser in place.

of *blue* atoms, and we know that they should belong to the distribution $P_0(\mathbf{R})$. In this case we again have an interference pattern. The which-way information is lost because, *with the shutters open*, the fact that the detector is still in its ground state does not allow us to find out where the photon is. Indeed, after many atoms are observed, we should see *red* and *blue* interference patterns. If these are *not* correlated with the observation of the detector, then they “lose their color” and we just get the sum, which is the pattern without interference.

In this thought experiment, one sees that which-way information can be obtained without taking into account any momentum transfer to the atoms that pass through the double slits. This happens because the path of the atom can be correlated with the behavior of a part of the apparatus with which the atom is entangled.

An experiment that follows in spirit, though not in detail, the proposal by Scully et al. has been carried out by S. Durr, T. Nonn, and G. Rempe, *Nature*, **395**, 33 (1998), and it bears out the quantum mechanical expectations.

Supplement 20-B

The Creation of GHZ States

The apparatus developed by Bouwmeester et al.¹ is shown in Fig. 20B-1. A short pulse of ultraviolet light passes through a nonlinear crystal, creating two pairs of photons, close enough in time that in terms of the counter time resolution they appear simultaneously. The photons move along paths a and b , and each pair of photons is entangled in that the polarization states are perpendicular to each other (the notation H and V is used for horizontal and vertical in the plane perpendicular to the propagation of the photons), in such a way that each pair may be described by the state

$$\frac{1}{\sqrt{2}} (|H\rangle_a |V\rangle_b - |V\rangle_a |H\rangle_b) \quad (20B-1)$$

The arm a leads to a polarizing beam-splitter. It acts to transmit H photons, which then continue to a detector, labeled T . This means that

$$|H\rangle_a \rightarrow |H\rangle_T \quad (20B-2)$$

The V photons are reflected. They move along the arm α and are made to pass through a $\lambda/2$ plate, which rotates their polarization (“ V ”) through 45° . At the polarizing beam splitter, the V -component is deflected to counter D_1 , while the H -component goes on to counter D_2 . This means that

$$|V\rangle_a \rightarrow \frac{1}{\sqrt{2}} (|V\rangle_1 + |H\rangle_2) \quad (20B-3)$$

The photons going along the arm b are directed to a polarization-independent beam-splitter, so that the photons reaching BS have a 50% chance of passing through to detector D_3 and a 50% chance of being deflected along the arm β . The photons moving along β strike the polarizing beam-splitter. The H photons go on to the detector D_1 , while the V photons that continue along the arm β and go on to the detector D_2 . This implies that

$$|H\rangle_b \rightarrow \frac{1}{\sqrt{2}} (|H\rangle_3 + |H\rangle_1) \quad (20B-4)$$

while

$$|V\rangle_b \rightarrow \frac{1}{\sqrt{2}} (|V\rangle_3 + |V\rangle_2) \quad (20B-5)$$

¹D. Bouwmeester, J-W. Pan, M. Daniell, H. Weinfurter, and A. Zeilinger, *Phys. Rev. Lett.* **82**, 1345 (1999).

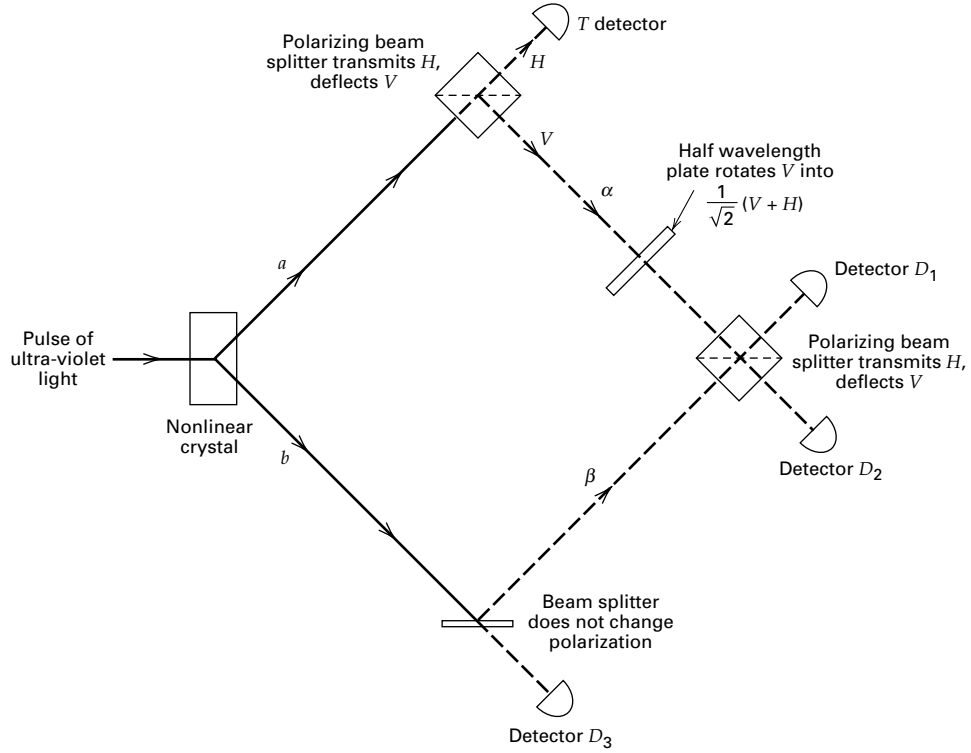


Figure 20B-1 Apparatus for the construction of GHZ states, based on the experiment of D. Bouwmeester, et al. *Phys. Rev. Letters*, **82**, 1345 (1999).

We may therefore see what happens to the entangled combination (20-1). We have

$$\begin{aligned} & \frac{1}{\sqrt{2}} (|H\rangle_a |V\rangle_b - |V\rangle_a |H\rangle_b) \rightarrow \\ & \frac{1}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} |H\rangle_T |V\rangle_3 + \frac{1}{\sqrt{2}} |H\rangle_T |V\rangle_2 \right] - \frac{1}{2} \frac{1}{\sqrt{2}} [(|V\rangle_1 |H\rangle_2)(|H\rangle_3 + |H\rangle_1)] = \quad (20B-6) \\ & = \frac{1}{2} \left[|H\rangle_T |V\rangle_3 + |H\rangle_T |V\rangle_2 - \frac{1}{\sqrt{2}} (|V\rangle_1 |H\rangle_3 + |V\rangle_1 |H\rangle_1 + |H\rangle_2 |H\rangle_1 + |H\rangle_2 |H\rangle_3) \right] \end{aligned}$$

We have a second photon pair, which has exactly the same form as (20B-6). If the second pair is emitted at a time such that it is possible to distinguish between the two pairs, then the form is that given in the last line of (20B-6), except that it is distinguished by a mark such as a prime. However, if the photons are emitted close enough in time so that the pairs cannot be distinguished, then we just take the last line of (20B-6) and multiply it by itself all over again. Although the product appears to have 36 terms, the experimental setup is such that all four counters click. This means that we have the following terms only:

$$- \frac{1}{2} \left[\frac{1}{\sqrt{2}} |H\rangle_T |V\rangle_3 |H\rangle_2 |H\rangle_1 + \frac{1}{\sqrt{2}} |H\rangle_T |V\rangle_2 |V\rangle_1 |H\rangle_3 \right]$$

so that the combination occurring in the counters complementary to T is

$$\frac{1}{\sqrt{2}} (|H\rangle_1 |H\rangle_2 |V\rangle_3 + |V\rangle_1 |V\rangle_2 |H\rangle_3)$$

This is a GHZ state. To make this look more like the GHZ state described in the text of the chapter, all we have to do is rotate the polarization detector D_3 so that $|H\rangle_3 \rightarrow |V\rangle_3$ and $|V\rangle_3 \rightarrow -|H\rangle_3$. The paper quoted above describes all the tests made to show that the state is indeed what it is expected to be. The experimental test showing that measurements on the GHZ state agree with the quantum mechanical predictions were carried out by the same authors, and the results can be found in *Nature* **403**, 515 (2000). To translate the algebra into the algebra of spin 1/2 states, we note that right- and left-circular polarization states are given by

$$|R\rangle = \frac{1}{\sqrt{2}}(|H\rangle - i|V\rangle)$$

$$|L\rangle = \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle)$$

The translation can now be carried out if we make the association

$$|R\rangle \rightarrow \chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |L\rangle \rightarrow \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

then

$$|H\rangle \rightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \quad |V\rangle \rightarrow \frac{i}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

which correspond to the eigenspinors of σ_x .

Supplement 20-C

The Density Operator

In all of our discussions we have dealt with the time development of physical systems, whose initial states were of the form

$$|\psi\rangle = \sum_n C_n |u_n\rangle \quad (20C-1)$$

Often such initial states are not the ones that are provided by the method of preparing the states. It may be that instead of a single ensemble consisting of identical states $|\psi\rangle$ we may be presented with a number of different ensembles on which measurements are to be performed. We may have a set of ensembles of the form

$$|\psi^{(i)}\rangle = \sum_n C_n^{(i)} |u_n\rangle \quad (20C-2)$$

and all we know is that the probability of finding an ensemble characterized by (i) is p_i , with

$$\sum_i p_i = 1 \quad (20C-3)$$

For example, we may have a beam of hydrogen atoms in an excited state, with fixed energy and orbital angular momentum l , but completely unpolarized, so that all m -values $-l \leq m \leq l$ are equally probable. In that case $p_m = 1/(2l + 1)$, independent of m . It is *not correct* to say that the beam is described by the wave function

$$|\psi\rangle = \sum_m C_m |Y_{lm}\rangle \quad (20C-4)$$

with $|C_m|^2 = 1/(2m + 1)$, since the physical situation represents $2m + 1$ independent beams, so that there is no phase relationship between different m -values.

The *density operator* formalism allows us to deal with both of these cases.

Pure State

Consider a pure state first. Define the density operator ρ by

$$\rho = |\psi\rangle\langle\psi| \quad (20C-5)$$

We can write this in the form

$$\rho = \sum_{m,n} C_n C_m^* |u_n\rangle\langle u_m| \quad (20C-6)$$

The matrix elements of ρ in the u_n basis are

$$\begin{aligned}\rho_{kl} &= \langle u_k | \rho | u_l \rangle = \langle u_k | \sum_{m,n} C_n C_m^* | u_n \rangle \langle u_m | u_l \rangle \\ &= C_k C_l^*\end{aligned}\quad (20C-7)$$

We observe that

(a)

$$\rho^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \rho \quad (20C-8)$$

(b)

$$\text{Tr } \rho = \sum_k \rho_{kk} = \sum_k |C_k|^2 = 1 \quad (20C-9)$$

(c) We can also write the expectation value of some observable as

$$\begin{aligned}\langle A \rangle &= \langle \psi | A | \psi \rangle = \sum_{m,n} C_m^* \langle u_m | A | u_n \rangle C_n \\ &= \sum_{m,n} C_m^* C_n A_{mn} = \sum_{m,n} A_{mn} \rho_{nm} \\ &= \text{Tr } (A\rho)\end{aligned}\quad (20C-10)$$

The results of equations (20C-8)–(20C-10) are independent of the choice of the complete set of basis vectors $|u_n\rangle$. To see this, consider the set $|v_n\rangle$. By the general expansion theorem, we can write

$$|v_n\rangle = \sum_m T_m^{(n)} |u_m\rangle$$

with

$$T_m^{(n)} = \langle u_m | v_n \rangle \equiv T_{mn}$$

Note that

$$\begin{aligned}\sum_n T_{mn} (T^+)_{nk} &= \sum_n T_{mn} T_{kn}^* = \sum_n \langle u_m | v_n \rangle \langle u_k | v_n \rangle^* \\ &= \sum_n \langle u_m | v_n \rangle \langle v_n | u_k \rangle = \delta_{mk}\end{aligned}$$

so that the matrix T is unitary. Then

$$\begin{aligned}|\psi\rangle &= \sum_k D_k |v_k\rangle \\ &= \sum_k D_k T_{kl} |u_l\rangle\end{aligned}$$

so that

$$C_l = D_k T_{kl} = (T^t)_{lk} D_k \equiv U_{lk} D_k$$

Since T is unitary, so is $U \equiv T^t$, the transpose of the matrix T . Thus

$$\rho_{kl} = C_k C_l^* = (U)_{km} D_m (U)_{ln}^* D_n^*$$

or

$$= (U \rho_D (U)^+)_{kl}$$

where ρ_D is the density operator in the v -basis. Thus

$$\rho_D = U^\dagger \rho U$$

It follows from the unitarity of U that the properties of ρ also apply to ρ_D .

Since $\rho = \rho^\dagger$, it follows that ρ may be diagonalized by a unitary transformation. This means that it is possible to choose a basis $|v_n\rangle$ such that ρ is diagonal. Since $\rho^2 = \rho$, this means that the eigenvalues can only be 1 and 0, and since $\text{Tr} \rho = 1$, only one eigenvalue can be 1, and all the others must be zero. Thus only one of the D_k can be nonvanishing. This means that in a suitably chosen basis, *a pure state is a state that is an eigenstate of a maximally commuting set of observables, whose eigenfunctions are the set $|v_n\rangle$.*

Mixed State

For a mixed state we define the density operator by

$$\rho = \sum_i |\psi^{(i)}\rangle p_i \langle \psi^{(i)}| \quad (20C-11)$$

In the $|u_n\rangle$ basis, this takes the form

$$\rho = \sum_{i,m,n} C_n^{(i)} C_m^{(i)*} p_i |u_n\rangle \langle u_m|$$

so that

$$\rho_{kl} = \langle u_k | \rho | u_l \rangle = \sum_i p_i C_k^{(i)} C_l^{(i)*} \quad (20C-12)$$

Note that $\rho_{kl} = \rho_{lk}^*$ so that ρ is hermitian. Since

$$\sum_n |C_n^{(i)}|^2 = 1$$

it follows that

$$\text{Tr} \rho = \sum_k \rho_{kk} = \sum_i p_i = 1 \quad (20C-13)$$

as before. Also

$$\begin{aligned} \langle A \rangle &= \sum_i p_i \langle \psi^{(i)} | A | \psi^{(i)} \rangle \\ &= \sum_i \sum_{mn} p_i \langle \psi^{(i)} | u_n \rangle \langle u_n | A | u_m \rangle \langle u_m | \psi^{(i)} \rangle \\ &= \sum_i \sum_{mn} p_i C_m^{(i)} C_n^{(i)*} A_{nm} \\ &= \sum_{mn} \rho_{mn} A_{nm} = \text{Tr} (\rho A) \end{aligned} \quad (20C-14)$$

as for pure state. On the other hand, it is no longer true that $\rho^2 = \rho$. In fact,

$$\rho^2 = \sum_j \sum_i |\psi^{(i)}\rangle p_i \langle \psi^{(i)} | \psi^{(j)}\rangle p_j \langle \psi^{(j)} | = \sum_i |\psi^{(i)}\rangle p_i^2 \langle \psi^{(i)} |$$

and, since $\sum_i p_i = 1$,

$$\text{Tr } \rho^2 = \sum_i p_i^2 < 1 \quad (20C-15)$$

for a mixed state.

It follows from the Schrödinger equation

$$\frac{d}{dt} |\psi^{(i)}\rangle = -\frac{i}{\hbar} H |\psi^{(i)}\rangle$$

and (since $H = H^\dagger$)

$$\frac{d}{dt} \langle \psi^{(i)} | = \frac{i}{\hbar} \langle \psi^{(i)} | H$$

that

$$\frac{d}{dt} \rho = -\frac{i}{\hbar} H \rho + \frac{i}{\hbar} \rho H = -\frac{i}{\hbar} [H, \rho] \quad (20C-16)$$

Note that the sign is opposite to the expression for the time rate of change of a general operator, which reads

$$\frac{d}{dt} A = \frac{i}{\hbar} [H, A]$$

The simplest application of the formalism is in the description of a beam of electrons or any other particles of spin 1/2. Here ρ is a 2×2 hermitian matrix. The most general form of such a matrix is

$$\rho = \frac{1}{2} (a \mathbf{1} + \mathbf{b} \cdot \boldsymbol{\sigma}) \quad (20C-17)$$

with a and $\mathbf{b}$ real. The condition $\text{tr } \rho = 1$ implies that $a = 1$. We can calculate ρ^2 as follows:

$$\begin{aligned} \rho^2 &= \frac{1}{4} (1 + \mathbf{b} \cdot \boldsymbol{\sigma})(1 + \mathbf{b} \cdot \boldsymbol{\sigma}) = \frac{1}{4} (1 + \mathbf{b}^2 + 2\mathbf{b} \cdot \boldsymbol{\sigma}) \\ &= \frac{1}{2} \left(\frac{1 + \mathbf{b}^2}{2} + \mathbf{b} \cdot \boldsymbol{\sigma} \right) \end{aligned} \quad (20C-18)$$

The density matrix ρ will describe a pure state only if $\rho^2 = \rho$ —that is, if $\mathbf{b}^2 = 1$. For a mixed state, it follows from (20C-15) that $\mathbf{b}^2 < 1$.

Let us now obtain a physical interpretation for $\mathbf{b}$. Consider a mixture of spin 1/2 beams. Each of the beams will have electrons aligned along either the z - or x - or y -axis. The fraction of particles that are in an eigenstate of σ_z with eigenvalue $+1$ will be denoted by $f_3^{(+)}$. The fractions that are in an eigenstate of σ_z with eigenvalue -1 will be denoted by $f_3^{(-)}$ and so on, so that

$$f_3^{(+)} + f_3^{(-)} + f_1^{(+)} + f_1^{(-)} + f_2^{(+)} + f_2^{(-)} = 1 \quad (20C-19)$$

The eigenstates, with eigenvalues ± 1 of σ_z , σ_x , and σ_y , are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \quad \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \end{pmatrix}$$

Thus the density matrix has the form

$$\begin{aligned}\rho = & f_3^{(+)} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + f_3^{(-)} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + f_1^{(+)} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\ & + f_1^{(-)} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} + f_2^{(+)} \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{pmatrix} \\ & + f_2^{(-)} \begin{pmatrix} 1/\sqrt{2} & i/\sqrt{2} \\ -i/\sqrt{2} & i/\sqrt{2} \end{pmatrix}\end{aligned}$$

A little algebra, with the use of (20C-19), shows that this can be written in the form

$$\rho = \frac{1}{2} + \frac{1}{2} \boldsymbol{\sigma} \cdot \mathbf{P} \quad (20C-20)$$

where $P_i = f_i^{(+)} - f_i^{(-)}$. The fraction of particles in a mixture that is aligned in the $+z$ -direction minus the fraction that is aligned in the $-z$ -direction is called the *polarization* in the z -direction, and we denote it by P_3 . Similarly for the other directions. Thus by comparing (20C-20) with (20C-18), we can interpret $\mathbf{b}$ as the net polarization vector $\mathbf{P}$ of the beam. In the case of beams of atoms of angular momentum l , the most general ρ is a $(2l+1) \times (2l+1)$ hermitian matrix, and the interpretation of the elements is more complicated. Further discussion of the density matrix is beyond the scope of this book.

Physical Constants¹

N_0 (Avogadro's number)	$6.0221367(36) \times 10^{23}/\text{gm-mole}$
c (speed of light—definition)	$2.99792458 \times 10^8 \text{ m/s}$
e (electron charge)	$1.60217733(49) \times 10^{-19} \text{ C}$
1 eV (unit of energy)	$1.60217733(49) \times 10^{-19} \text{ J}$
$\hbar$ (Planck constant/ 2π)	$1.05457266(63) \times 10^{-34} \text{ J} \cdot \text{s}$
	$6.5821220(20) \times 10^{-22} \text{ MeV} \cdot \text{s}$
$\alpha = e^2/4\pi\epsilon_0\hbar c$ (fine structure constant)	$1/137.0359895(61)$
k (Boltzmann constant)	$1.380658(12) \times 10^{-23} \text{ J/K}$
m_e (electron mass)	$9.1093897(54) \times 10^{-31} \text{ kg}$
	$0.51099906(15) \text{ MeV}/c^2$
m_p (proton mass)	$1.6726231(10) \times 10^{-27} \text{ kg}$
	$938.27231(28) \text{ MeV}/c^2$
m_n (neutron mass)	$1.6749286(1) \times 10^{-27} \text{ kg}$
	$939.56563(28) \text{ MeV}/c^2$
1 a.m.u. (m ($\text{C}^{12}/12$))	$1.6605402(10) \times 10^{-27} \text{ kg}$
	$931.49432(28) \text{ MeV}/c^2$
$a_0 = \hbar/m_e c \alpha$ (Bohr radius)	$0.529177249(24) \times 10^{-10} \text{ m}$
$R_0 = m_e c^2 \alpha^2/2$ (Rydberg)	$13.0605698(40) \text{ eV}$
G (gravitational constant)	$6.67259(85) \times 10^{-11} \text{ m}^3/\text{kg} \cdot \text{s}^2$
$\mu_B = e\hbar/2m_e$ (Bohr magneton)	$0.578838263(52) \times 10^{-10} \text{ MeV/T}$

¹The numbers are taken from a very interesting article by E. R. Cohen and B. N. Taylor in *Physics Today* (August 1994).

References¹

I. Books on Various Topics in Modern Physics²

- J. Bernstein, P. M. Fishbane, and S. Gasiorowicz, *Modern Physics*, Prentice Hall, Upper Saddle River, N.J., (2000).
- J. Brehm and W. J. Mullin, *Introduction to the Structure of Matter*, John Wiley, New York, 1989.
- R. Eisberg and R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles* (2nd edition), John Wiley, New York, 1985.

II. Textbooks and Monographs

- G. Baym, *Lectures on Quantum Mechanics*, W. A. Benjamin, New York, 1969.

This is a very appealing book, with just the right mixture of formalism, intuitive arguments, and applications. It should be considered an advanced book, accessible to the student who has covered the material in this book.

- H. A. Bethe and R. W. Jackiw, *Intermediate Quantum Mechanics* (2nd edition), W. A. Benjamin, New York, 1968.

This book contains detailed discussions of calculational methods applicable to the theory of atomic structure, multiplet splittings, the photoelectric effect, and atomic collisions. Much of the material is not to be found in any other textbook. The book is thus an advanced text, as well as an exhaustive reference book.

- H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, Springer-Verlag, Berlin/New York, 1957.

This reprint of the authors' article in the *Handbuch der Physik* is an elaborate, detailed, definitive treatment of the problem at hand. It is a book about atoms and not about quantum mechanics, and the level is high. It is an excellent reference book.

- D. Bohm, *Quantum Theory*, Dover, New York, 1989.

The book is discursively written, on a level comparable to the present book. The author pays

much attention to the principles of quantum theory and gives an excellent discussion of the quantum theory of the measurement process. There are few applications and not many problems.

- S. Borowitz, *Fundamentals of Quantum Mechanics*, W. A. Benjamin, New York, 1967.

This is a well-written book, about half of which is devoted to the theory of waves and to classical mechanics. The level is comparable to that of the present book.

- C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics*, John Wiley, New York, 1977.

This is an encyclopedic book of more than a thousand pages. The student will find detailed coverage of many aspects of atomic physics. The mathematical level is quite a bit above that of *Quantum Physics*.

- R. H. Dicke and J. P. Wittke, *Introduction to Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1960.

I enjoyed this book very much. It is on a level comparable to the present book, and discusses a few topics, notably quantum statistics, that are not treated here. The problems are excellent.

- P. A. M. Dirac, *The Principles of Quantum Mechanics* (4th edition), Oxford University Press (Clarendon), Oxford, England, 1958.

This is a superb book by one of the major creators of quantum mechanics. The student who has studied the material in this book will have no trouble with Dirac; if at all serious about mastering quantum mechanics, he or she should sooner or later go through Dirac's book.

- R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.

In 1948, R. P. Feynman proposed a different formulation of quantum mechanics. In this book the equivalence of this formulation to the standard theory is demonstrated, and the "path integral" expression for the general amplitude is exploited in a number of calculations. The selection of material is very interesting, and the point of view is different from the one developed by the author. Thus this

¹There are an ever growing number of textbooks on quantum mechanics. I have studied from some of them and read some others on specific subjects. The list is limited, and no book is criticized by its omission from the list.

²These books are at a level that is somewhat lower than that of *Quantum Physics*.

somewhat more advanced book presents an excellent complement to this book.

R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Vol. 3, *Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1965.

In this introduction to quantum mechanics, Feynman abandons the path integral and approaches the subject from the point of view of state vectors. A large number of fascinating examples are discussed with the minimum of formal apparatus. A superb complementary book, whose only shortcoming is the absence of problems.

K. Gottfried, *Quantum Mechanics*, Vol. 1, *Fundamentals*, W. A. Benjamin, New York, 1966.

This is a very advanced book, distinguished by the care with which the various topics are discussed. The treatment of the measurement process and of invariance principles is excellent. The student who has mastered the material in this book should be able to read Gottfried's book, provided he or she has acquired the necessary mathematical equipment.

D. Griffiths, *Introduction to Quantum Mechanics*, Prentice Hall, Englewood Cliffs, N.J., 1995.

This well-written, attractive book is roughly on the level of Dicke and Wittke or Saxon. It contains a nice selection of topics, including a discussion of the geometric phase.

G. Greenstein and A. G. Zajonc, *The Quantum Challenge*, Jones and Bartlett, Sudbury, Mass., 1997.

This is a fine book that deals with the foundations of quantum mechanics from the point of view of basic experiments. It is a real pleasure to read and makes few mathematical demands on the reader.

W. Heisenberg, *The Physical Principles of the Quantum Theory*, Dover, New York, 1930.

This reprint of some 1930 lectures given by Heisenberg on the physical significance of the quantum theory still makes good reading. The discussion of the uncertainty relations is particularly useful.

T. E. Jordan, *Quantum Mechanics in Simple Matrix Form*, John Wiley, New York, 1986.

This is a very interesting book in that it deals with quantum mechanics in the form originally created by W. Heisenberg, M. Born, and P. Jordan. Calculations became much simpler with the (equivalent) wave mechanics approach of Schrödinger, but as this brief book shows, one can do a lot with simple matrices. This is the place where one can find the matrix solution to the hydrogen atom, and there is a nice discussion of Bell's inequalities.

H. A. Kramers, *Quantum Mechanics*, Interscience, New York, 1957.

This book by one of the founders of the subject is at its best in the discussion of spin and the introduction to relativistic quantum theory, both rather advanced subjects. The student who is comfortable with quantum mechanics will find browsing through this book enjoyable and rewarding.

L. D. Landau and E. M. Lifshitz, *Quantum Mechanics (Nonrelativistic Theory)* (2nd edition), Addison-Wesley, Reading, Mass., 1965.

The book by Landau and Lifshitz is one of a series of superb books covering all of theoretical physics. It is hard to think of this as a textbook for any but the most sophisticated students. Any student, however, once he reaches the advanced level, will find much that is useful in this book. There is an assumed mathematical facility on the part of the student.

R. L. Liboff, *Introductory Quantum Mechanics*, Addison Wesley Longman, Reading, Mass., 1998.

This is a very attractive book, written on roughly the same mathematical level as *Quantum Physics*. It is an excellent alternative reference, with a more detailed treatment of solid state physics applications.

H. J. Lipkin, *Quantum Mechanics—New Approaches to Selected Topics*, North-Holland, Amsterdam, 1973.

Lipkin's book deals with a number of advanced topics in the application of quantum mechanics, in a simple way. The physics is always at the forefront of the discussion, and a student who has a good command of *Quantum Physics* will get much benefit and pleasure out of this book.

A. Messiah, *Quantum Mechanics* (in 2 volumes), John Wiley, New York, 1968.

This book gives a complete coverage of quantum theory from the treatment of one-dimensional potentials through the quantization of the electromagnetic field and the relativistic wave equation of Dirac. It is an advanced book, and it assumes a mathematical sophistication that few first-year graduate students possess. It is an extremely worthwhile book.

E. Merzbacher, *Quantum Mechanics* (3rd edition), John Wiley, New York, 1998.

Together with the books of Schiff and Sakurai, this is the standard first-year graduate text in quantum mechanics. Recent editions have enlarged it, but the standard of economy and taste have been maintained.

R. Omnes, *Understanding Quantum Mechanics*, Princeton University Press, Princeton, N.J., 1999.

This excellent book covers a great deal of the work done in recent years on extending or modifying the

generally accepted approach to the meaning of quantum mechanics.

H. C. Ohanian, *Principles of Quantum Mechanics*, Prentice Hall, Englewood Cliffs, N.J., 1990.

The basics of quantum mechanics are covered at more or less the same level as the present book. There is a nice discussion of Bell's theorems and a brief overview of some of the concerns about "the interpretation of quantum mechanics."

D. Park, *Introduction to the Quantum Theory*, (3rd edition) McGraw Hill, New York, 1992.

This excellent book is written at a level comparable to *Quantum Physics* and complements it nicely. A variety of topics not covered in this book are treated, such as the formation of cloud chamber tracks, Bell's inequalities, and more details on the motion of electrons in a periodic lattice.

W. Pauli, *Die Allgemeinen Prinzipien der Wellenmechanik, Handbuch der Physik*, Vol. 5/1, Springer-Verlag, Berlin/New York, 1958.

The advanced student who reads German will find in this reprint of a 1930 article by Pauli a concise definitive discussion of quantum mechanics. There are no applications, but all of the important matters are there.

P. J. E. Peebles, *Quantum Mechanics*, Princeton University Press, Princeton, N.J., 1992.

This is a nicely written textbook at the undergraduate level. It differs from other books at this level (except that by Bohm) by a detailed discussion of what is really meant by the measurement process in quantum mechanics.

A. B. Pippard, *The Physics of Vibration*, Cambridge University Press, Cambridge, England, 1978.

This book covers all kinds of oscillators, classically and quantum mechanically. It is not a textbook in the usual sense. It is a delight to read.

J. L. Powell and B. Crasemann, *Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1961.

The strength of this book is in the painstaking working out of all of the mathematical details of wave mechanics and matrix mechanics. Probably all of the mathematical aspects of these subjects that have been bypassed in this book can be found here. There is a good discussion of the WKB approximation and of the general properties of second-order differential equations. There are relatively few applications, and there are more exercises than problems.

R. W. Robinett, *Quantum Mechanics*, Oxford University Press, Oxford, England, 1997.

This book is more or less on the same level as *Quantum Physics* and it is useful as an alternative

reference. There is a nice treatment of two-dimensional quantum mechanics.

J. Schwinger, *Quantum Mechanics*, Springer Verlag, Berlin and Heidelberg, 2001.

The lectures of J. Schwinger on quantum mechanics were edited posthumously by B-G. Englert. They provide an original approach to the formalism, and they show clearly Schwinger's wizardry in mathematical physics.

J. J. Sakurai, *Modern Quantum Mechanics* (S. F. Tuan, Editor), Addison-Wesley, Reading, Mass., 1994.

This excellent book by the late J. J. Sakurai is written at a somewhat more advanced level than *Quantum Physics*, like the books by Merzbacher and Schiff. This first-year graduate textbook really does have a modern flavor with a selection of topics that leads quite naturally into the areas of advanced quantum mechanics of interest to particle physicists. The book has an excellent collection of problems.

D. S. Saxon, *Elementary Quantum Mechanics*, Holden-Day, San Francisco, 1968.

This book is on the same level as the present one, and it is a useful reference, since the selection of topics is just a little different, as is the emphasis and the choice of applications. The book contains an excellent set of problems.

L. I. Schiff, *Quantum Mechanics* (3rd edition), McGraw-Hill, New York, 1968.

This is one of the standard first-year graduate textbooks. It is perhaps a little too compact, and thus most suitable for the well-prepared student. The level of mathematical sophistication assumed is above that of the reader of the present book.

F. Schwabl, *Quantum Mechanics*, Springer-Verlag, New York, 1992.

This is a more advanced book with an interesting selection of topics.

R. Shankar, *Principles of Quantum Mechanics*, Plenum, New York, 1980.

This is a more sophisticated and mathematically advanced book than the present book. It contains alternative treatment of some of the topics discussed in this book, and is a good reference.

M. P. Silverman, *And Yet it Moves: Strange Systems and Subtle Questions in Physics*, Cambridge University Press, New York, 1993.

This book deals in a qualitative way with a number of topics that involve the basic principles of quantum physics. It is fun to read and quite accessible to the student who has covered a good part of *Quantum Physics*.

Index

Note: Entries listed as Supplements are available at the Wiley web site: www.wiley.com/college/gasiorowicz

- Addition of spin and orbital angular momentum, 171
- Addition of two angular momenta, 171
 - details (Supplement 10-A), W-43
- Addition of two spins, 166
- Adiabatic theorem (problem), 244
- Adiabatic transition rate, 238
- Aharonov-Bohm effect, 258
 - details (Supplement 16-A), W-75
 - flux quantization (Supplement 16-A), W-79
- Alpha decay (Supplement 4-B), W-16
- Ammonia molecule, 80
- Angular momentum
 - commutation relations of operators, 121
 - eigenvalues for L^2 and L_z , 123
 - eigenfunctions in spherical coordinates, 124
 - in spherical coordinates (Supplement 7-B), W-30
 - matrix representation, 150
 - raising and lowering operators, 112
- Anomalous Zeeman effect, 191
- Antisymmetrization of wave function, 203
 - When is antisymmetrization needed?, 204
- Associated Laguerre polynomials, 138
- Associated Legendre polynomials (Supplement 7-B), W-34
- Asymptotic series in Stark effect, 185
- Autoionization, 223
- Balmer formula, Bohr derivation, 16
- Band structure (Supplement 4-C), W-21
- Barrier penetration, 73
- Bell's inequalities (footnote), 324
- Bessel functions (Supplement 16-B), W-80
- Blackbody radiation, 1
 - Planck distribution, 3
 - details (Supplement 1-A), W-2
 - quanta of energy, 3
 - Rayleigh-Jeans distribution, 2
 - Stefan-Boltzmann law, 3
- Bloch's theorem (Supplement 4-C), W-20
- Bohr postulates of atomic structure, 15
- Born approximation in scattering, 302
- Bosons, 202
- Bras (*see* Dirac notation)
- Breit-Wigner formula for resonant scattering, 297
- Building-up principle (Supplement 14-B), W-55
- Classical limit (Ehrenfest theorem), 105
 - charge in constant magnetic field, 252
 - for angular momentum, 128
- Clebsch-Gordan coefficients (Supplement 10-A), W-46
- Coherent states (problem), 118
- Cold emission, 73
- Collision cross section, 287
- Completeness relation, 97, 109
- Compton effect, 7
- Conservation of total momentum (Supplement 13-A), W-49
- Cooling of atoms, 273
- Correspondence principle, 20
 - dipole transition rate, 268
- Coulomb scattering in Born approximation, 304
- Cyclotron frequency, 163
- Dark states (Supplement 18-B), W-89
- De Broglie wavelength, 10
- Decoherence, 327
- Degeneracy, 59
- Degeneracy and simultaneous observables, 100
- Degeneracy pressure, 211
- Degenerate perturbation theory, 177–180
- Degenerate stars, 212
- Delayed-choice experiment, 319
- Delta functions (Supplement 2-A), W-7
- Delta function potential
 - bound states, 81
 - double well bound states, 81
 - matching condition, 67
- Density operator (Supplement 20-C), W-99
- Detailed balance (problem), 315
- Diagonalization of matrices, 148
- Differential cross section, 290
- Dirac notation, 107
 - scalar product, 108
 - connection with wave function, 110
- Eigenvalues and eigenfunctions, 45
- Einstein A and B coefficients (Supplement 1-A), W-1
- Electric dipole approximation, 263
- Electric dipole selection rules, 265
- Electric quadrupole matrix element, 265
- Electromagnetic field
 - energy density, 261
 - gauge transformations, 247, 249, 257
 - Hamiltonian, 248
 - Maxwell's equations, 246
 - magnetic dipole term, 265
 - quantization (Supplement 18-A), W-83
 - semiclassical treatment, 261
- Electromagnetically induced transparency (Supplement 18-B), W-89

- Electromagnetic potentials, 247
- Electron diffraction, 10
- Energy conservation in transitions, 240
- Energy shift in perturbation theory, 175
- Entanglement, 170, 318, 322
- Exclusion principle (*see* Pauli exclusion principle)
- Expansion coefficients (interpretation), 53, 97
- Expansion postulate, 52
- Expectation values, 36
- Exponential decay formula (Supplement 15A)
 - and energy uncertainty relation, 243
- Fermi energy, 209
- Fermions, 202
- Ferromagnetism an exchange force, 223
- Feynman-Hellmann theorem (problem), 234
- Flux in one-dimensional motion, 66
- Flux of probability current, 35
- Free particle in spherical coordinates, 140
- GHZ state creation (Supplement 20-B), W-95
- GHZ test of “realism,” 324
- Gauge transformations, 247
- Hamiltonian operator, 94
- Harmonic oscillator
 - eigenfunctions, eigenvalues, 87
 - matrix representation, 147
 - raising and lowering operators, 113
 - solution by operator methods, 112
 - zero point energy, 8
- Harmonic time variation of potential, 238
- Hartree approximation (Supplement 14-A), W-51
- Heisenberg picture, 116
- Heisenberg uncertainty relations, 32
- Helium atom, 215
 - effect of exclusion principle on spectrum, 217
 - effect of electron-electron repulsion, 218
 - exchange interaction, 220
- Hermitian operators, 38, 96, 99
- Hermite polynomials
 - general properties, 100
- Hund’s spectroscopic rules, 222
 - Supplement 14-B, W-58
- Hydrogen atom, 133
 - anomalous Zeeman effect, 191
 - degeneracy of spectrum, 135
 - energy spectrum, 135
- H_2^+ molecule, 228
- Hydrogen molecule (Supplement 14-C), W-64
 - hyperfine structure, 194
 - radial eigenfunctions, 138
 - radial equation, 134
 - reduced mass effect, 135, 196
 - relativistic kinetic energy shift, 188
 - spin-orbit coupling, 188
- Identical particles, 201
 - scattering cross section, 305
- Independent electron approximation for atoms
 - 226
- Induced emission, 270
- Inelastic scattering, 291
- Infinite well, 47
 - spherical infinite well solutions, 143
- Interaction picture (footnote), 237
- Interaction picture (Supplement 15-B), W-73
- Internal conversion (problem), 245
- Kapitza-Dirac effect, 13
- Kets (*see* Dirac notation)
- Kronig-Penney model (Supplement 4-C), W-21
- Landau levels, 253
- Lasers, 270
- Legendre polynomials, 126
- Legendre polynomials (Supplement 7B), W-33
- Lenz vector (footnote), 137
- Levi-Civita symbol (Supplement 10-B), W-47
- Levinson’s theorem, 299
- Lifetime, 243
 - Weisskopf-Wigner (Supplement 15-A), W-69
- Linewidth (Supplement 15-A), W-71
- Matrices in quantum mechanics, 146
- Matrix elements defined, 148
- Matrix elements of angular momentum operators,
 - 150
- Maxwell’s equations, 246
 - different formulation (Supplement 10-B), W-47
- Molecular orbitals, 229
 - in Supplement 14-C, W-65
- Molecular spectra, 231
- Momentum eigenfunctions, 56
 - normalization, 57
- Momentum operator, 37
- Momentum-space wave function, 38, 96
- Mossbauer effect, 282
- Neutron stars, 211
- One-dimensional double well, 79
- One-dimensional potential barrier, 71
- One-dimensional potential step, 66
- One-dimensional potential well
 - bound states, 75
 - reflection and transmission, 66
- Optical theorem, 291
- Orbitals, 222
- Orthogonality of eigenfunctions
 - general, 95
 - harmonic oscillator, 88
 - hermitian operators, 99, 109
- Paramagnetic resonance, 164
- Parity, 60
 - and Pauli principle, 206
 - of pion, 207

- Pauli exclusion principle, 202
 - and filling of levels, 208
 - astrophysical implications, 211
 - effect on molecular spectra, 232
 - singlet and triplet spin states, 206
- Pauli spin matrices, 159
- Pauli theorem (Supplement 8-A), W-35
- Periodic boundary conditions, 241
- Periodic potentials (Supplement 4-C), W-19
- Periodic table (Supplement 14-B), W-61
- Perturbation theory
 - energy shifts, 174, 176
 - time-dependent, 236
 - time-independent, 174
- Phase shift in scattering (Supplement 8-B), W-39
- Phase space, 241
- Photodisintegration of deuteron, 266
 - (problem), 314
- Photoelectric effect, 5
 - cross section, 308
- Planck distribution (*see* Blackbody radiation)
- Plane wave in terms of spherical harmonics
 - (Supplement 8-B), W-41
- Polarization, 261
- Potential step, 66
- Probability, survey (Supplement 2-B), W-11
- Probability current (*see* Flux of probability current)
- Projection operators, 111
- Quantum eraser, 320
- Quantum eraser (Supplement 20-A), W-92
- Quantum Hall effect, 255
- Quantum jumps with single atoms, 279
- Quantum measurement, 325
- Rabi frequency, 278
- Radiative decays, 260
- Ramsauer-Townsend Effect (*see* Transmission resonance)
- Recoilless emission (*see* Mossbauer effect)
- Reduced mass effect, 19, 196
- Resonant scattering, 297
- Resonant scattering (Supplement 15-A), W-72
- Ritz variational principle, 224
 - application to helium ground state energy, 225
- Rotating wave approximation, 165, 276
- Rotational invariance (Supplement 7-A), W-28
- Rutherford planetary model of the atom, 15
 - lifetime (Supplement 1-B), W-4
- Scanning tunneling microscope (STM), 74
- Scattering
 - black disk, 292
 - Born approximation, 302
 - by atoms on a lattice, 307
 - connection with binding energy, 299
 - effective range formula, 300
 - identical particles, 305
 - low energy, 294
 - resonant, 296
- Scattering matrix (problems), 90, 91
- Schrödinger equation
 - central potential, 132
 - connection with operator formulation, 114
 - for hydrogen atom, 133
 - in three dimensions, 131
 - time dependent, 30
 - time-independent, 45
- Schrödinger picture, 116
- Selection rules
 - in anomalous Zeeman effect, 193
 - in radiative decays, 264
- Shadow scattering, 293
- Shell structure, 227
 - atomic (Supplement 14-B), W-55
 - infinite well in three dimensions, 144
- Simultaneous observables, 100
- Slater determinant, 204
- Spherical Bessel functions, 141
- Spherical harmonics
 - completeness relation, 128
 - functional form (table), 127
 - probability distributions, 141
- Spin-orbit coupling, 188
- Spin 1/2 operators in matrix form, 158
 - eigenstates, 159
 - entanglement of two spin states, 170
 - intensity rules (Supplement 17-A), W-81
 - intrinsic magnetic moment, 162
 - singlet and triplet eigenstates, 167
 - spin addition, 166
 - spin precession, 163
 - spin-dependent potentials, 169
 - spin-dependent scattering, 300
 - square well *S*-wave scattering, 298
- Square well continuum solutions (Supplement 8-B), W-38
- Stark effect, 180
- Stern-Gerlach experiment, 19
- Thomas precession effect, 189
- Thomas-Reiche-Kuhn sum rule (problem), 186
- Three-level system, 278
- Three-level systems (Supplement 18-B), W-86
- Time dependence of expectation values, 103
- Time dependence of operators, 116
- Time-dependent perturbation theory, 236
- Trace of matrix, 149
- Transition rate, 240, 260
 - for many particle decays, 242
 - $2p \rightarrow 1s$ transition, 267
- Transmission resonance, 70
- Tunneling in nuclear physics (Supplement 4-B), W-16
- Tunneling in one-dimensional potential barrier, 72

336 Index

- Twenty-one centimeter line, 196
- Two-particle system, 199
- Uncertainty relations, 32
 - definition of uncertainty, 102
 - illustrations, 32
- Valence bonds (Supplement 14-C), W-66
- Valence electrons (Supplement 14-C), W-67
- Variational principle (*see* Ritz variational principle)
- Vector spaces, 98
- Virial theorem (problem), 93, 145
- WKB approximation (Supplement 4-A), W-14
- Wave function
 - momentum space, 38, 56, 96
 - probability interpretation, 28, 34
- Wave packets, 25
 - motion and spreading of wave packets, 27
- Wave-particle duality, 23
- Which-way measurements (Supplement 20-A), W-91
- Zeeman effect, 191
- Zero-zero transitions, 267