

# QUANTUM MECHANICS

# QUANTUM MECHANICS

---

## A Conceptual Approach

HENDRIK F. HAMEKA

 **WILEY-INTERSCIENCE**

A John Wiley & Sons, Inc. Publication

Copyright © 2004 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.  
Published simultaneously in Canada.

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 Section 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, Inc., 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400, fax 978-646-8600, or on the web at [www.copyright.com](http://www.copyright.com). 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.

**Limit of Liability/Disclaimer of Warranty:** While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. at 317-572-3993 or fax 317-572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print, however, may not be available in electronic format.

***Library of Congress Cataloging-in-Publication Data:***

Hameka, Hendrik F.

Quantum mechanics : a conceptual approach / Hendrik F. Hameka.

p. cm.

Includes index.

ISBN 0-471-64965-1 (pbk. : acid-free paper)

1. Quantum theory. I. Title.

QC174.12.H353 2004

530.12-dc22

2004000645

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

*To Charlotte*

# CONTENTS

<b>Preface</b>	<b>xi</b>
<b>1 The Discovery of Quantum Mechanics</b>	<b>1</b>
I Introduction, 1	
II Planck and Quantization, 3	
III Bohr and the Hydrogen Atom, 7	
IV Matrix Mechanics, 11	
V The Uncertainty Relations, 13	
VI Wave Mechanics, 14	
VII The Final Touches of Quantum Mechanics, 20	
VIII Concluding Remarks, 22	
<b>2 The Mathematics of Quantum Mechanics</b>	<b>23</b>
I Introduction, 23	
II Differential Equations, 24	
III Kummer's Function, 25	
IV Matrices, 27	
V Permutations, 30	
VI Determinants, 31	

VII	Properties of Determinants, 32	
VIII	Linear Equations and Eigenvalues, 35	
IX	Problems, 37	
<b>3</b>	<b>Classical Mechanics</b>	<b>39</b>
I	Introduction, 39	
II	Vectors and Vector Fields, 40	
III	Hamiltonian Mechanics, 43	
IV	The Classical Harmonic Oscillator, 44	
V	Angular Momentum, 45	
VI	Polar Coordinates, 49	
VII	Problems, 51	
<b>4</b>	<b>Wave Mechanics of a Free Particle</b>	<b>52</b>
I	Introduction, 52	
II	The Mathematics of Plane Waves, 53	
III	The Schrödinger Equation of a Free Particle, 54	
IV	The Interpretation of the Wave Function, 56	
V	Wave Packets, 58	
VI	Concluding Remarks, 62	
VII	Problems, 63	
<b>5</b>	<b>The Schrödinger Equation</b>	<b>64</b>
I	Introduction, 64	
II	Operators, 66	
III	The Particle in a Box, 68	
IV	Concluding Remarks, 71	
V	Problems, 72	
<b>6</b>	<b>Applications</b>	<b>73</b>
I	Introduction, 73	
II	A Particle in a Finite Box, 74	

III	Tunneling, 78	
IV	The Harmonic Oscillator, 81	
V	Problems, 87	
<b>7</b>	<b>Angular Momentum</b>	<b>88</b>
I	Introduction, 88	
II	Commuting Operators, 89	
III	Commutation Relations of the Angular Momentum, 90	
IV	The Rigid Rotor, 91	
V	Eigenfunctions of the Angular Momentum, 93	
VI	Concluding Remarks, 96	
VII	Problems, 96	
<b>8</b>	<b>The Hydrogen Atom</b>	<b>98</b>
I	Introduction, 98	
II	Solving the Schrödinger Equation, 99	
III	Deriving the Energy Eigenvalues, 101	
IV	The Behavior of the Eigenfunctions, 103	
V	Problems, 106	
<b>9</b>	<b>Approximate Methods</b>	<b>108</b>
I	Introduction, 108	
II	The Variational Principle, 109	
III	Applications of the Variational Principle, 111	
IV	Perturbation Theory for a Nondegenerate State, 113	
V	The Stark Effect of the Hydrogen Atom, 116	
VI	Perturbation Theory for Degenerate States, 119	
VII	Concluding Remarks, 120	
VIII	Problems, 120	
<b>10</b>	<b>The Helium Atom</b>	<b>122</b>
I	Introduction, 122	

II	Experimental Developments, 123	
III	Pauli's Exclusion Principle, 126	
IV	The Discovery of the Electron Spin, 127	
V	The Mathematical Description of the Electron Spin, 129	
VI	The Exclusion Principle Revisited, 132	
VII	Two-Electron Systems, 133	
VIII	The Helium Atom, 135	
IX	The Helium Atom Orbitals, 138	
X	Concluding Remarks, 139	
XI	Problems, 140	
<b>11</b>	<b>Atomic Structure</b>	<b>142</b>
I	Introduction, 142	
II	Atomic and Molecular Wave Function, 145	
III	The Hartree-Fock Method, 146	
IV	Slater Orbitals, 152	
V	Multiplet Theory, 154	
VI	Concluding Remarks, 158	
VII	Problems, 158	
<b>12</b>	<b>Molecular Structure</b>	<b>160</b>
I	Introduction, 160	
II	The Born-Oppenheimer Approximation, 161	
III	Nuclear Motion of Diatomic Molecules, 164	
IV	The Hydrogen Molecular Ion, 169	
V	The Hydrogen Molecule, 173	
VI	The Chemical Bond, 176	
VII	The Structures of Some Simple Polyatomic Molecules, 179	
VIII	The Hückel Molecular Orbital Method, 183	
IX	Problems, 189	
	<b>Index</b>	<b>191</b>



# PREFACE

The physical laws and mathematical structure that constitute the basis of quantum mechanics were derived by physicists, but subsequent applications became of interest not just to the physicists but also to chemists, biologists, medical scientists, engineers, and philosophers. Quantum mechanical descriptions of atomic and molecular structure are now taught in freshman chemistry and even in some high school chemistry courses. Sophisticated computer programs are routinely used for predicting the structures and geometries of large organic molecules or for the identification and evaluation of new medicinal drugs. Engineers have incorporated the quantum mechanical tunneling effect into the design of new electronic devices, and philosophers have studied the consequences of some of the novel concepts of quantum mechanics. They have also compared the relative merits of different axiomatic approaches to the subject.

In view of the widespread applications of quantum mechanics to these areas there are now many people who want to learn more about the subject. They may, of course, try to read one of the many quantum textbooks that have been written, but almost all of these textbooks assume that their readers have an extensive background in physics and mathematics; very few of these books make an effort to explain the subject in simple non-mathematical terms.

In this book we try to present the fundamentals and some simple applications of quantum mechanics by emphasizing the basic concepts and by keeping the mathematics as simple as possible. We do assume that the reader is familiar with elementary calculus; it is after all not possible to explain the Schrödinger equation to someone who does not know what a derivative or an integral is. Some of the mathematical techniques that are essential for understanding quantum mechanics, such as matrices and determinants, differential equations, Fourier analysis, and so on are

described in a simple manner. We also present some applications to atomic and molecular structure that constitute the basis of the various molecular structure computer programs, but we do not attempt to describe the computation techniques in detail.

Many authors present quantum mechanics by means of the axiomatic approach, which leads to a rigorous mathematical representation of the subject. However, in some instances it is not easy for an average reader to even understand the axioms, let alone the theorems that are derived from them. I have always looked upon quantum mechanics as a conglomerate of revolutionary new concepts rather than as a rigid mathematical discipline. I also feel that the reader might get a better understanding and appreciation of these concepts if the reader is familiar with the background and the personalities of the scientists who conceived them and with the reasoning and arguments that led to their conception. Our approach to the presentation of quantum mechanics may then be called historic or conceptual but is perhaps best described as pragmatic. Also, the inclusion of some historical background makes the book more readable.

I did not give a detailed description of the various sources I used in writing the historical sections of the book because many of the facts that are presented were derived from multiple sources. Some of the material was derived from personal conversations with many scientists and from articles in various journals. The most reliable sources are the original publications where the new quantum mechanical ideas were first proposed. These are readily available in the scientific literature, and I was intrigued in reading some of the original papers. I also read various biographies and autobiographies. I found Moore's biography of Schrödinger, Constance Reid's biographies of Hilbert and Courant, Abraham Pais' reminiscences, and the autobiographies of Elsasser and Casimir particularly interesting. I should mention that Kramers was the professor of theoretical physics when I was a student at Leiden University. He died before I finished my studies and I never worked under his supervision, but I did learn quantum mechanics by reading his book and by attending his lectures.

Finally I wish to express my thanks to Mrs. Alice Chen for her valuable help in typing and preparing the manuscript.

HENDRIK F. HAMEKA

---

# 1

---

## THE DISCOVERY OF QUANTUM MECHANICS

### I. INTRODUCTION

The laws of classical mechanics were summarized in 1686 by *Isaac Newton* (1642–1727) in his famous book *Philosophiae Naturalis Principia Mathematica*. During the following 200 years, they were universally used for the theoretical interpretation of all known phenomena in physics and astronomy. However, towards the end of the nineteenth century, new discoveries related to the electronic structure of atoms and molecules and to the nature of light could no longer be interpreted by means of the classical Newtonian laws of mechanics. It therefore became necessary to develop a new and different type of mechanics in order to explain these newly discovered phenomena. This new branch of theoretical physics became known as quantum mechanics or wave mechanics.

Initially quantum mechanics was studied solely by theoretical physicists or chemists, and the writers of textbooks assumed that their readers had a thorough knowledge of physics and mathematics. In recent times the applications of quantum mechanics have expanded dramatically. We feel that there is an increasing number of students who would like to learn the general concepts and fundamental features of quantum mechanics without having to invest an excessive amount of time and effort. The present book is intended for this audience.

We plan to explain quantum mechanics from a historical perspective rather than by means of the more common axiomatic approach. Most fundamental concepts of quantum mechanics are far from self-evident, and they gained general

**TABLE 1-1. Pioneers of Quantum Mechanics**


---

Niels Henrik David Bohr (1885–1962)
Max Born (1882–1970)
Louis Victor Pierre Raymond, Duc de Broglie (1892–1989)
Pieter Josephus Wilhelmus Debye (1884–1966)
Paul Adrien Maurice Dirac (1902–1984)
Paul Ehrenfest (1880–1933)
Albert Einstein (1879–1955)
Samuel Abraham Goudsmit (1902–1978)
Werner Karl Heisenberg (1901–1976)
David Hilbert (1862–1943)
Hendrik Anton Kramers (1894–1952)
Wolfgang Ernst Pauli (1900–1958)
Max Karl Ernst Ludwig Planck (1858–1947)
Erwin Rudolf Josef Alexander Schrödinger (1887–1961)
Arnold Johannes Wilhelm Sommerfeld (1868–1951)
George Eugene Uhlenbeck (1900–1988)

---

acceptance only because there were no reasonable alternatives for the interpretation of new experimental discoveries. We believe therefore that they may be easier to understand by learning the motivation and the line of reasoning that led to their discovery.

The discovery of quantum mechanics makes an interesting story, and it has been the subject of a number of historical studies. It extended over a period of about 30 years, from 1900 to about 1930. The historians have even defined a specific date, namely, December 14, 1900, as the birth date of quantum mechanics. On that date the concept of quantization was formulated for the first time.

The scientists who made significant contributions to the development of quantum mechanics are listed in Table 1.1. We have included one mathematician in our list, namely, David Hilbert, a mathematics professor at Göttingen University in Germany, who is often regarded as the greatest mathematician of his time. Some of the mathematical techniques that were essential for the development of quantum mechanics belonged to relatively obscure mathematical disciplines that were known only to a small group of pure mathematicians. Hilbert had studied and contributed to these branches of mathematics, and he included the material in his lectures. He was always available for personal advice with regard to mathematical problems, and some of the important advances in quantum mechanics were the direct result of discussions with Hilbert. Eventually his lectures were recorded, edited, and published in book form by one of his assistants, *Richard Courant* (1888–1972). The book, *Methods of Mathematical Physics*, by R. Courant and D. Hilbert, was published in 1924, and by a happy coincidence it contained most of the mathematics that was important for the study and understanding of quantum mechanics. The book became an essential aid for most physicists.

Richard Courant was a famous mathematician in his own right. He became a colleague of Hilbert's as a professor of mathematics in Göttingen, and he was instrumental in establishing the mathematical institute there. In spite of his accomplishments, he was one of the first Jewish professors in Germany to be dismissed from his position when the Nazi regime came to power (together with Max Born, who was a physics professor in Göttingen). In some respects Courant was fortunate to be one of the first to lose his job because at that time it was still possible to leave Germany. He moved to New York City and joined the faculty of New York University, where he founded a second institute of mathematics. Born was also able to leave Germany, and he found a position at Edinburgh University.

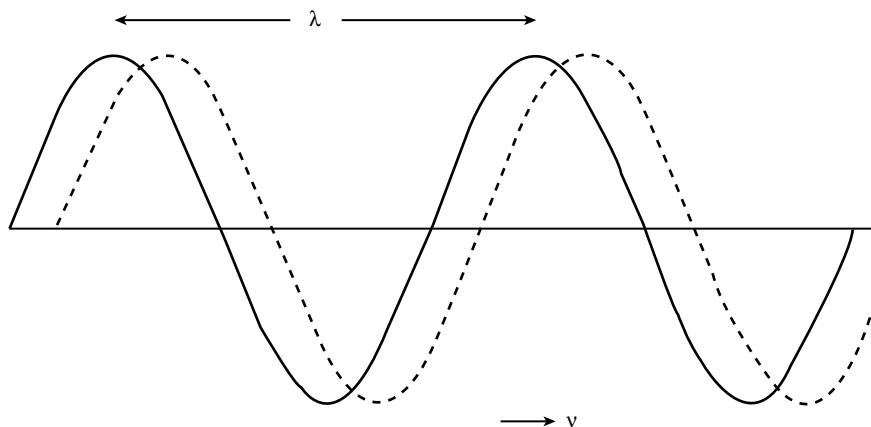
It may be of interest to mention some of the interpersonal relations between the physicists listed in Table 1-1. Born was Hilbert's first assistant and Sommerfeld was Klein's mathematics assistant in Göttingen. After Born was appointed a professor in Göttingen, his first assistants were Pauli and Heisenberg. Debye was Sommerfeld's assistant in Aachen and when the latter became a physics professor in Munich, Debye moved with him to Munich. Kramers was Bohr's first assistant in Copenhagen, and he succeeded Ehrenfest as a physics professor in Leiden. Uhlenbeck and Goudsmit were Ehrenfest's students. We can see that the physicists lived in a small world, and that they all knew each other.

In this chapter, we present the major concepts of quantum mechanics by giving a brief description of the historical developments leading to their discovery. In order to explain the differences between quantum mechanics and classical physics, we outline some relevant aspects of the latter in Chapter 3. Some mathematical topics that are useful for understanding the subject are presented in Chapter 2. In subsequent chapters, we treat various simple applications of quantum mechanics that are of general interest. We attempt to present the material in the simplest possible way, but quantum mechanics involves a fair number of mathematical derivations. Therefore, by necessity, some mathematics is included in this book.

## II. PLANCK AND QUANTIZATION

The introduction of the revolutionary new concept of quantization was a consequence of Planck's efforts to interpret experimental results related to black body radiation. This phenomenon involves the interaction between heat and light, and it attracted a great deal of attention in the latter part of the nineteenth century.

We have all experienced the warming effect of bright sunlight, especially when we wear dark clothing. The sunlight is absorbed by our dark clothes, and its energy is converted to heat. The opposite effect may be observed when we turn on the heating element of an electric heater or a kitchen stove. When the heating element becomes hot it begins to emit light, changing from red to white. Here the electric energy is first converted to heat, which in turn is partially converted to light.



**Figure 1-1** Sketch of a one-dimensional wave.

It was found that the system that was best suited for quantitative studies of the interaction between light and heat was a closed container since all the light within the vessel was in equilibrium with its walls. The light within such a closed system was referred to as black body radiation. It was, of course, necessary to punch a small hole in one of the walls of the container in order to study the characteristics of the black body radiation. One interesting finding of these studies was that these characteristics are not dependent on the nature of the walls of the vessel.

We will explain in Chapter 4 that light is a wavelike phenomenon. A wave is described by three parameters: its propagation velocity  $v$ ; its wavelength  $\lambda$ , which measures the distance between successive peaks; and its frequency  $\nu$  (see Figure 1-1). The frequency is defined as the inverse of the period  $T$ , that is, the time it takes the wave to travel a distance  $\lambda$ . We have thus

$$v = \lambda/T = \lambda\nu \quad (1-1)$$

White light is a composite of light of many colors, but monochromatic light consists of light of only one color. The color of light is determined solely by its frequency, and monochromatic light is therefore light with a specific characteristic frequency  $\nu$ . All different types of light waves have the same propagation velocity  $c$ , and the frequency  $\nu$  and wavelength  $\lambda$  of a monochromatic light wave are therefore related as

$$c = \lambda\nu \quad (1-2)$$

It follows that a monochromatic light wave has both a specific frequency  $\nu$  and a specific wavelength  $\lambda$ .

The experimentalists were interested in measuring the energy of black body radiation as a function of the frequency of its components and of temperature. As more experimental data became available, attempts were made to represent these data by empirical formulas. This led to an interesting controversy because it turned out that one formula, proposed by *Wilhelm Wien* (1864–1928), gave an accurate representation of the high-frequency data, while another formula, first proposed by *John William Strutt, Lord Rayleigh* (1842–1919), gave an equally good representation of the low-frequency results. Unfortunately, these two formulas were quite different, and it was not clear how they could be reconciled with each other.

Towards the end of the nineteenth century, a number of theoreticians attempted to find an analytic expression that would describe black body radiation over the entire frequency range. The problem was solved by Max Planck, who was a professor of theoretical physics at the University of Berlin at the time. Planck used thermodynamics to derive a formula that coincided with Wien's expression for high frequencies and with Rayleigh's expression for low frequencies. He presented his result on October 19, 1900, in a communication to the German Physical Society. It became eventually known as Planck's radiation law.

Even though Planck had obtained the correct theoretical expression for the temperature and frequency dependence of black body radiation, he was not satisfied. He realized that his derivation depended on a thermodynamic interpolation formula that, in his own words, was nothing but a lucky guess.

Planck decided to approach the problem from an entirely different direction, namely, by using a statistical mechanics approach. Statistical mechanics was a branch of theoretical physics that described the behavior of systems containing large numbers of particles and that had been developed by *Ludwig Boltzmann* (1844–1906) using classical mechanics.

In applying Boltzmann's statistical methods, Planck introduced the assumption that the energy  $E$  of light with frequency  $\nu$  must consist of an integral number of energy elements  $\epsilon$ . The energy  $E$  was therefore quantized, which means that it could change only in a discontinuous manner by an amount  $\epsilon$  that constituted the smallest possible energy element occurring in nature. We are reminded here of atomic theory, in which the atom is the smallest possible amount of matter. By comparison, the energy quantum is the smallest possible amount of energy. We may also remind the reader that the concept of quantization is not uncommon in everyday life. At a typical auction the bidding is quantized since the bids may increase only by discrete amounts. Even the Internal Revenue Service makes use of the concept of quantization since our taxes must be paid in integral numbers of dollars, the financial quanta.

Planck's energy elements became known as quanta, and Planck even managed to assign a quantitative value to them. In order to analyze the experimental data of black body radiation, Planck had previously introduced a new fundamental constant to which he assigned a value of  $6.55 \times 10^{-27}$  erg sec. This constant is now known as Planck's constant and is universally denoted by the symbol  $h$ . Planck proposed that the magnitude of his energy elements or quanta was given by

$$\epsilon = h\nu \tag{1-3}$$

Many years later, in 1926, the American chemist *Gilbert Newton Lewis* (1875–1946) introduced the now common term photon to describe the light quanta.

Planck reported his analysis at the meeting of the German Physical Society on December 14, 1900, where he read a paper entitled “On the Theory of the Energy Distribution Law in the ‘Normalspectrum.’” This is the date that historians often refer to as the birth date of quantum mechanics.

Privately Planck believed that he had made a discovery comparable in importance to Newton’s discovery of the laws of classical mechanics. His assessment was correct, but during the following years his work was largely ignored by his peers and by the general public.

We can think of a number of reasons for this initial lack of recognition. The first and obvious reason was that Planck’s paper was hard to understand because it contained a sophisticated mathematical treatment of an abstruse physical phenomenon. A second reason was that his analysis was not entirely consistent even though the inconsistencies were not obvious. However, the most serious problem was that Planck was still too accustomed to classical physics to extend the quantization concept to its logical destination, namely, the radiation itself. Instead Planck introduced a number of electric oscillators on the walls of the vessel, and he assumed that these oscillators were responsible for generating the light within the container. He then applied quantization to the oscillators or, at a later stage, to the energy transfer between the oscillators and the radiation. This model added unnecessary complications to his analysis.

Einstein was aware of the inconsistencies of Planck’s theory, but he also recognized the importance of its key feature, the concept of quantization. In 1905 he proposed that this concept should be extended to the radiation field itself. According to Einstein, the energy of a beam of light was the sum of its light quanta  $h\nu$ . In the case of monochromatic light, these light quanta or photons all have the same frequency and energy, but in the more general case of white light they may have different frequencies and a range of energy values.

Einstein used these ideas to propose a theoretical explanation of the photoelectric effect. Two prominent physicists, *Joseph John Thomson* (1856–1940) and *Philipp Lenard* (1862–1947), discovered independently in 1899 that electrons could be ejected from a metal surface by irradiating the surface with light. They found that the photoelectric effect was observed only if the frequency of the incident light was above a certain threshold value  $\nu_0$ . When that condition is met, the velocity of the ejected electrons depends on the frequency of the incident light but not on its intensity, while the number of ejected electrons depends on the intensity of the light but not on its frequency.

Einstein offered a simple explanation of the photoelectric effect based on the assumption that the incident light consisted of the light quanta  $h\nu$ . Let us further suppose that the energy required to eject one electron is defined as  $eW$ , where  $e$  is the electron charge. It follows that only photons with energy in excess of  $eW$  are capable of ejecting electrons; consequently

$$h\nu_0 = eW \tag{1-4}$$



A photon with a frequency larger than  $\nu_0$  has sufficient energy to eject an electron and, its energy surplus  $E$

$$E = h\nu - eW \quad (1-5)$$

is equal to the kinetic energy of the electron. The number of ejected electrons is, of course, determined by the number of light quanta with frequencies in excess of  $\nu_0$ . In this way, all features of the photoelectric effect were explained by Einstein in a simple and straightforward manner. Einstein's theory was confirmed by a number of careful experiments during the following decade. It is interesting to note that the threshold frequency  $\nu_0$  corresponds to ultraviolet light for most metals but to visible light for the alkali metals (e.g., green for sodium). The excellent agreement between Einstein's equation and the experimental data gave strong support to the validity of the quantization concept as applied to the radiation field.

The idea became even more firmly established when it was extended to other areas of physics. The specific heat of solids was described by the rule of Dulong and Petit, which states that the molar specific heats of all solids have the same temperature-independent value. This rule was in excellent agreement with experimental bindings as long as the measurements could not be extended much below room temperature. At the turn of the twentieth century, new techniques were developed for the liquefaction of gases that led to the production of liquid air and, subsequently, liquid hydrogen and helium. As a result, specific heats could be measured at much lower temperatures, even as low as a few degrees above the absolute temperature minimum. In this way, it was discovered that the specific heat of solids decreases dramatically with decreasing temperature. It even appears to approach zero when the temperature approaches its absolute minimum.

The law of Dulong and Petit had been derived by utilizing classical physics, but it soon became clear that the laws of classical physics could not account for the behavior of specific heat at lower temperatures. It was Einstein who showed in 1907 that the application of the quantization concept explained the decrease in specific heat at lower temperatures. A subsequent more precise treatment by Debije produced a more accurate prediction of the temperature dependence of the specific heat in excellent agreement with experimental bindings.

Since the quantization concept led to a number of successful theoretical predictions, it became generally accepted. It played an important role in the next advance in the development of quantum mechanics, which was the result of problems related to the study of atomic structure.

### III. BOHR AND THE HYDROGEN ATOM

Atoms are too small to be studied directly, and until 1900 much of the knowledge of atomic structure had been obtained indirectly. Spectroscopic measurements made significant contributions in this respect.

An emission spectrum may be observed by sending an electric discharge through a gas in a glass container. This usually leads to dissociation of the gas molecules.

The atoms then emit the energy that they have acquired in the form of light of various frequencies. The emission spectrum corresponds to the frequency distribution of the emitted light.

It was discovered that most atomic emission spectra consist of a number of so-called spectral lines; that is, the emitted light contains a number of specific discrete frequencies. These frequencies could be measured with a high degree of accuracy.

The emission frequencies of the hydrogen atom were of particular interest. The four spectral lines in the visible part of the spectrum were measured in 1869 by the Swedish physicist *Anders Jöns Ångström* (1814–1874). It is interesting to note that the unit of length that is now commonly used for the wavelength of light is named after him. The Ångström unit (symbol Å) is defined as  $10^{-8}$  cm. The wavelengths of the visible part of the spectrum range from 4000 to 8000 Å.

The publication of Ångström's highly precise measurements stimulated some interest in detecting a relationship between those numbers. In 1885 the Swiss physics teacher *Johann Jakob Balmer* (1825–1898) made the surprising discovery that the four wavelengths measured by Ångström could be represented exactly by the formula

$$\lambda = \frac{Am^2}{m^2 - 4} \quad m = 3, 4, 5, 6 \quad (1-6)$$

A few years later, *Johannes Robert Rydberg* (1854–1919) proposed a more general formula

$$\nu = R \left( \frac{1}{n^2} - \frac{1}{m^2} \right) \quad n, m = 1, 2, 3 \quad (m > n) \quad (1-7)$$

which accurately represented all frequencies of the hydrogen emission spectrum, including those outside the visible part of the spectrum;  $R$  became known as the Rydberg constant.

It was perceived first by Rydberg and later by *Walter Ritz* (1878–1909) that Eq. (1-7) is a special case of a more general formula that is applicable to the spectral frequencies of atoms in general. It is known as the combination principle, and it states that all the spectral frequencies of a given atom are differences of a much smaller set of quantities, defined as terms

$$\nu = |T_i - T_j| \quad (1-8)$$

We should understand that 10 terms determine 45 frequencies, 100 terms 4950 frequencies, and so on.

The above rules were, of course, quite interesting, and there was no doubt about their validity since they agreed with the experimental spectral frequencies to the many decimal points to which the latter could be measured. At the same time, there was not even the remotest possibility that they could be explained on the basis of the laws of classical physics and mechanics.

Meanwhile, a great deal of information about the structure of atoms had become available through other experiments. During a lecture on April 30, 1897 at the Royal Institution in Great Britain, *Joseph John Thomson* (1856–1940) first proposed the existence of subatomic particles having a negative electric charge and a mass considerably smaller than that of a typical atomic mass. The existence of these particles was confirmed by subsequent experiments, and they became known by the previously proposed name electrons.

Thomson's discovery of the electron was followed by a large number of experimental studies related to atomic structure. We will not describe these various discoveries in detail; suffice it to say that in May 1911 they helped *Ernest Rutherford* (1871–1937) propose a theoretical model for the structure of the atom that even today is generally accepted.

According to Rutherford, an atom consists of a nucleus with a radius of approximately  $3 \times 10^{-12}$  cm, having a positive electric charge, surrounded by a number of electrons with negative electric charges at distances of the order of  $1 \text{ \AA}$  ( $10^{-8}$  cm) from the central nucleus. The simplest atom is hydrogen, where one single electron moves in an orbit around a much heavier nucleus.

Rutherford's atomic model has often been compared to our solar system. In a similar way, we may compare the motion of the electron around its nucleus in the hydrogen atom to the motion of the moon around the Earth. There are, however, important differences between the two systems. The moon is electrically neutral, and it is kept in orbit by the gravitational attraction of the Earth. It also has a constant energy since outside forces due to the other planets are negligible. The electron, on the other hand, has an electric charge, and it dissipates energy when it moves. According to the laws of classical physics, the energy of the electron should decrease as a function of time. In other words, the assumption of a stable electronic orbit with constant energy is inconsistent with the laws of classical physics. Since classical physics could not explain the nature of atomic spectra, the scientists were forced to realize that the laws of classical physics had lost their universal validity, and that they ought to be reconsidered and possibly revised.

The dilemma was solved by Niels Bohr, who joined Rutherford's research group in Manchester in 1912 after a short and unsatisfactory stay in Thomson's laboratory in Cambridge. Bohr set out to interpret the spectrum of the hydrogen atom, but in the process he made a number of bold assumptions that were developed into new fundamental laws of physics. His first postulate assumed the existence of a discrete set of stationary states with constant energy. A system in such a stationary state neither emits nor absorbs energy.

It may be interesting to quote Bohr's own words from a memoir he published in 1918:

- I. That an atomic system can, and can only, exist permanently in a certain series of states corresponding to a discontinuous series of values for its energy, and that consequently 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 states. These states will be denoted as the "stationary states" of the system.

II. That the radiation absorbed or emitted during a transition between two stationary states is “unifrequent” and possesses a frequency  $\nu$  given by the relation

$$E' - E'' = h\nu$$

where  $h$  is Planck’s constant and where  $E'$  and  $E''$  are the values of the energy in the two states under consideration.

The second part of Bohr’s statement refers to his second postulate, which states that a spectroscopic transition always involves two stationary states; it corresponds to a change from one stationary state to another. The frequency  $\nu$  of the emitted or absorbed radiation is determined by Planck’s relation  $\Delta E = h\nu$ . This second postulate seems quite obvious today, but it was considered revolutionary at the time.

Bohr successfully applied his theory to a calculation of the hydrogen atom spectrum. An important result was the evaluation of the Rydberg constant. The excellent agreement between Bohr’s result and the experimental value confirmed the validity of both Bohr’s theory and Rutherford’s atomic model.

Bohr’s hydrogen atom calculation utilized an additional quantum assumption, namely, the quantization of the angular momentum, which subsequently became an important feature of quantum mechanics. It should be noted here that Ehrenfest had in fact proposed this same correct quantization rule for the angular momentum a short time earlier in 1913. The rule was later generalized by Sommerfeld.

In the following years, Bohr introduced a third postulate that became known as the correspondence principle. In simplified form, this principle requires that the predictions of quantum mechanics for large quantum numbers approach those of classical mechanics.

Bohr returned to Copenhagen in 1916 to become a professor of theoretical physics. In that year Kramers volunteered to work with him, and Bohr was able to offer him a position as his assistant. Kramers worked with Bohr until 1926, when he was appointed to the chair of theoretical physics at the University of Utrecht in the Netherlands. Meanwhile, Bohr helped to raise funds for the establishment of an Institute for Theoretical Physics. He was always stimulated by discussions and personal interactions with other physicists, and he wanted to be able to accommodate visiting scientists and students. The Institute for Theoretical Physics was opened in 1921 with Bohr as its first director. During the first 10 years of its existence, it attracted over sixty visitors and became an international center for the study of quantum mechanics.

In spite of its early successes, the old quantum theory as it was practiced in Copenhagen between 1921 and 1925 left much to be desired. It gave an accurate description of the hydrogen atom spectrum, but attempts to extend the theory to larger atoms or molecules had little success. A much more serious shortcoming of the old quantum theory was its lack of a logical foundation. In its applications to atoms or molecules, random and often arbitrary quantization rules were introduced after the system was described by means of classical electromagnetic theory. Many physicists felt that there was no fundamental justification for these quantization rules other than the fact that they led to correct answers.

The situation improved significantly during 1925 and 1926 due to some dramatic advances in the theory that transformed quantum mechanics from a random set of rules into a logically consistent scientific discipline. It should be noted that most of the physicists listed in Table 1-1 contributed to these developments.

#### IV. MATRIX MECHANICS

During 1925 and 1926 two different mathematical descriptions of quantum mechanics were proposed. The first model became known as matrix mechanics, and its initial discovery is attributed to Heisenberg. The second model is based on a differential equation proposed by Schrödinger that is known as the Schrödinger equation. It was subsequently shown that the two different mathematical models are equivalent because they may be transformed into one another. The discovery of matrix mechanics preceded that of the Schrödinger equation by about a year, and we discuss it first.

Matrix mechanics was first proposed in 1925 by Werner Heisenberg, who was a 23-year-old graduate student at the time. Heisenberg began to study theoretical physics with Sommerfeld in Munich. He transferred to Göttingen to continue his physics study with Born when Sommerfeld temporarily left Munich to spend a sabbatical leave in the United States. After receiving his doctoral degree, Heisenberg joined Bohr and Kramers in Copenhagen. He became a professor of theoretical physics at Leipzig University, and he was the recipient of the 1932 Nobel Prize in physics at the age of 31.

Heisenberg felt that the quantum mechanical description of atomic systems should be based on physical observable quantities only. Consequently, the classical orbits and momenta of the electrons within the atom should not be used in a theoretical description because they cannot be observed. The theory should instead be based on experimental data that can be derived from atomic spectra. Each line in an atomic spectrum is determined by its frequency  $\nu$  and by its intensity. The latter is related to another physical observable known as its transition moment. A typical spectral transition between two stationary states  $n$  and  $m$  is therefore determined by the frequency  $\nu(n, m)$  and by the transition moment  $x(n, m)$ . Heisenberg now proposed a mathematical model in which physical quantities could be presented by sets that contained the transition moments  $x(n, m)$  in addition to time-dependent frequency terms. When Heisenberg showed his work to his professor, Max Born, the latter soon recognized that Heisenberg's sets were actually matrices, hence the name matrix mechanics.

We present a brief outline of linear algebra, the theory of matrices and determinants in Chapter 2. Nowadays linear algebra is the subject of college mathematics courses taught at the freshman or sophomore level, but in 1925 it was an obscure branch of mathematics unknown to physicists. However, by a fortunate coincidence, linear algebra was the subject of the first chapter in the newly published book *Methods of Mathematical Physics* by Courant and Hilbert. *Ernst Pascual Jordan* (1902–1980) was Courant's assistant who helped write the chapter on

matrices, and he joined Born and Heisenberg in deriving the rigorous formulation of matrix mechanics. The results were published in a number of papers by Born, Jordan, and Heisenberg, and the discovery of matrix mechanics is credited to these three physicists.

We do not give a detailed description of matrix mechanics because it is rather cumbersome, but we attempt to outline some of its main features. In the classical description, the motion of a single particle of mass  $m$  is determined by its position coordinates  $(x, y, z)$  and by the components of its momentum  $(p_x, p_y, p_z)$ . The latter are defined as the products of the mass  $m$  of the particle and its velocity components  $(v_x, v_y, v_z)$ :

$$p_x = mv_x, \text{ etc.} \quad (1-9)$$

Here  $p_x$  is called conjugate to the coordinate  $x$ ,  $p_y$  to  $y$ , and  $p_z$  to  $z$ . The above description may be generalized to a many-particle system by introducing a set of generalized coordinates  $q_i$  and conjugate moments  $p_i$ . These generalized coordinates and momenta constitute the basis for the formulation of matrix mechanics.

In Chapter 2 we discuss the multiplication rules for matrices, and we will see that the product  $\mathbf{A} \cdot \mathbf{B}$  of two matrices that we symbolically represent by the bold-face symbols  $\mathbf{A}$  and  $\mathbf{B}$  is not necessarily equal to the product  $\mathbf{B} \cdot \mathbf{A}$ . In matrix mechanics the coordinates  $q_i$  and moments  $p_i$  are symbolically represented by matrices. For simplicity, we consider one-dimensional motion only. The quantization rule requires that the difference between the two matrix products  $\mathbf{p} \cdot \mathbf{q}$  and  $\mathbf{q} \cdot \mathbf{p}$  be equal to the identity matrix  $\mathbf{I}$  multiplied by a factor  $h/2\pi$ . Since the latter combination occurred frequently, a new symbol  $\hbar$  was introduced by defining

$$\hbar = h/2\pi \quad (1-10)$$

The quantization rule could therefore be written as

$$\mathbf{p} \cdot \mathbf{q} - \mathbf{q} \cdot \mathbf{p} = \hbar \cdot \mathbf{I} \quad (1-11)$$

In order to determine the stationary states of the system, it is first necessary to express the energy of the system as a function of the coordinate  $q$  and the momentum  $p$ . This function is known as the Hamiltonian function  $H$  of the system, and it is defined in Section 3.III. The matrix  $\mathbf{H}$  representing the Hamiltonian is obtained by substituting the matrices  $q$  and  $p$  into the analytical expression for the Hamiltonian.

The stationary states of the system are now derived by identifying expressions for the matrix representations  $\mathbf{q}$  and  $\mathbf{p}$  that lead to a diagonal form for  $\mathbf{H}$ —in other words, to a matrix  $\mathbf{H}$  where all nondiagonal elements are zero. The procedure is well defined, logical, and consistent, and it was successfully applied to derive the stationary states of the harmonic oscillator. However, the mathematics that is required for applications to other systems is extremely cumbersome, and the practical use of matrix mechanics was therefore quite limited.

There is an interesting and amusing anecdote related to the discovery of matrix mechanics. When Heisenberg first showed his work to Born, he did not know what matrices were and Born did not remember very much about them either, even though he had learned some linear algebra as a student. It was therefore only natural that they turned to Hilbert for help. During their meeting, Hilbert mentioned, among other things, that matrices played a role in deriving the solutions of differential equations with boundary conditions. It was this particular feature that was later used to prove the equivalence of matrix mechanics and Schrödinger's differential equation. Later on, Hilbert told some of his friends laughingly that Born and Heisenberg could have discovered Schrödinger's equation earlier if they had just paid more attention to what he was telling them. Whether this is true or not, it makes a good story. It is, of course, true that Schrödinger's equation is much easier to use than matrix mechanics.

## V. THE UNCERTAINTY RELATIONS

Heisenberg's work on matrix mechanics was of a highly specialized nature, but his subsequent formulation of the uncertainty relations had a much wider appeal. They became known outside the scientific community because no scientific background is required to understand or appreciate them.

It is well known that any measurement is subject to a margin of error. Even though the accuracy of experimental techniques has been improved in the course of time and the possible errors of experimental results have become smaller, they are still of a finite nature. Classical physics is nevertheless designed for idealized situations based on the assumption that it is in principle possible to have exact knowledge of a system. It is then also possible to derive exact predictions about the future behavior of the system.

Heisenberg was the first to question this basic assumption of classical physics. He published a paper in 1927 where he presented a detailed new analysis of the nature of experimentation. The most important feature of his paper was the observation that it is not possible to obtain information about the nature of a system without causing a change in the system. In other words, it may be possible to obtain detailed information about a system through experimentation, but as a result of this experimentation, it is no longer the same system and our information does not apply to the original system. If, on the other hand, we want to leave the system unchanged, we should not disturb it by experimentation. Heisenberg's observation became popularly known as the uncertainty principle; it is also referred to as the indeterminacy principle.

Heisenberg summarized his observation at the conclusion of his paper as follows: "In the classical law 'if we know the presence exactly we can predict the future exactly' it is the assumption and not the conclusion that is incorrect."

A second feature of Heisenberg's paper dealt with the simultaneous measurement of the position or coordinate  $q_i$  of a particle and of its conjugate momentum  $p_i$ . If, for example, we consider one-dimensional motion, it should be clear that we

must monitor the motion of a particle over a certain distance  $\Delta q$  in order to determine its velocity  $v$  and momentum  $p$ . It follows that the uncertainty  $\Delta p$  in the result of the momentum measurement is inversely related to the magnitude  $\Delta q$ ; the larger  $\Delta q$  is the smaller  $\Delta p$  is, and vice versa. Heisenberg now proposed that there should be a lower limit for the product of  $\Delta q$  and  $\Delta p$  and that the magnitude of this lower limit should be consistent with the quantization rule (1-11) of matrix mechanics. The result is

$$\Delta q \cdot \Delta p > \hbar \quad (1-12)$$

Heisenberg proposed a similar inequality for the uncertainty  $\Delta E$  in measurements of the energy of the system during a time interval  $\Delta t$ :

$$\Delta E \cdot \Delta t > \hbar \quad (1-13)$$

It should again be obvious that the accuracy of energy measurements should improve if more time is available for the experiment. The quantitative magnitude of the lower limit of the product of  $\Delta E$  and  $\Delta t$  is consistent with the quantization rules of matrix mechanics.

In Section 4.V, of Chapter 4, we describe a special situation that was created by Heisenberg himself where the product of  $\Delta q$  and  $\Delta p$  is equal to  $\hbar/2$ , exactly half of the value of the uncertainty relation (1-12). However, this is an idealized special case, and it does not invalidate the principle of the uncertainty relations.

Heisenberg's work became of interest not only to physicists but also to philosophers because it led to a reevaluation of the ideas concerning the process of measurement and to the relations between theory and experiment. We will not pursue these various ramifications.

## VI. WAVE MECHANICS

We have already mentioned that the formulation of wave mechanics was the next important advance in the formulation of quantum mechanics. In this section we give a brief description of the various events that led to its discovery, with particular emphasis on the contributions of two scientists, Louis de Broglie and Erwin Schrödinger.

Louis de Broglie was a member of an old and distinguished French noble family. The family name is still pronounced as "breuil" since it originated in Piedmonte. The family includes a number of prominent politicians and military heroes; two of the latter were awarded the title "Marshal of France" in recognition of their outstanding military leadership. One of the main squares in Strasbourg, the Place de Broglie, and a street in Paris are named after family members.

Louis de Broglie was educated at the Sorbonne in Paris. Initially he was interested in literature and history, and at age 18 he graduated with an arts degree. However, he had developed an interest in mathematics and physics, and he decided to



pursue the study of theoretical physics. He was awarded a second degree in science in 1913, but his subsequent physics studies were then interrupted by the First World War. He was fortunate to be assigned to the army radiotelegraphy section at the Eiffel Tower for the duration of the war. Because of this assignment, he acquired a great deal of practical experience working with electromagnetic radio waves.

In 1920 Louis resumed his physics studies. He again lived in the family mansion in Paris, where his oldest brother, *Maurice, Duc de Broglie* (1875–1960), had established a private physics laboratory. Maurice was a prominent and highly regarded experimental physicist, and at the time he was interested in studying the properties of X rays. It is not surprising that the two brothers, Louis and Maurice, developed a common interest in the properties of X rays and had numerous discussions on the subject.

Radio waves, light waves, and X rays may all be regarded as electromagnetic waves. The various waves all have the same velocity of wave propagation  $c$ , which is considered a fundamental constant of nature and which is roughly equal to 300,000 km/sec. The differences between the types of electromagnetic waves are attributed to differences in wavelength. Visible light has a wavelength of about 5000 Å, whereas radio waves have much longer wavelengths of the order of 100 m and X rays have much shorter wavelengths of the order of 1 Å. The relation between velocity of propagation  $c$ , wavelength  $\lambda$ , and frequency  $\nu$  is in all cases given by Eq. (1-2).

When Louis de Broglie resumed his physics studies in 1920, he became interested in the problems related to the nature of matter and radiation that arose as a result of Planck's introduction of the quantization concept. De Broglie felt that if light is emitted in quanta, it should have a corpuscular structure once it has been emitted. Nevertheless, most of the experimental information on the nature of light could be interpreted only on the basis of the wave theory of light that had been introduced by the Dutch scientist *Christiaan Huygens* (1629–1695) in his book *Traité de la Lumière . . .*, published in 1690.

The situation changed in 1922, when experimental work by the American physicist *Arthur Holly Compton* (1892–1962) on X ray scattering produced convincing evidence for the corpuscular nature of radiation. Compton measured the scattering of so-called hard X rays (of very short wavelengths) by substances with low atomic numbers—for instance, graphite. Compton found that the scattered X rays have wavelengths larger than the wavelength of the incident radiation and that the increase in wavelength is dependent on the scattering angle.

Compton explained his experimental results by using classical mechanics and by describing the scattering as a collision between an incident X ray quantum, assumed to be a particle, and an electron. The energy  $E$  and momentum  $p$  of the incident X ray quantum are assumed to be given by

$$E = h\nu \quad p = h\nu/c = h/\lambda \quad (1-14)$$

The energy and momentum of the electron before the collision are much smaller than the corresponding energy  $E$  and momentum  $p$  of the X ray quantum, and

they are assumed to be negligible. Compton found that there was perfect agreement between his calculations and the experimental results. Maurice de Broglie quickly became aware of what is now popularly known as the Compton effect.

In a later memoir Louis remarked that in conversations with his brother, they concluded that X rays could be regarded both as particles and as waves. In his Nobel Prize lecture, Louis de Broglie explained that he felt that it was necessary to combine the corpuscular and wave models and to assume that the motion of an X ray quantum particle is associated with a wave. Since the corpuscular and wave motions cannot be independent, it should be possible to determine the relation between the two concepts.

Louis de Broglie's hypothesis assumed that the motion of an X ray was of a corpuscular nature and that the particle motion was accompanied by a wave. The relation between particle and wave motion was described by Eq. (1-14). De Broglie now proceeded to a revolutionary extension of this idea, namely, that the model was not confined to X rays and other forms of radiation but that it should be applicable to all other forms of motion, in particular the motion of electrons. Consequently, a beam of electrons moving with momentum  $p$  should be associated with a wave with wavelength

$$\lambda = h/p \quad (1-15)$$

De Broglie supported his proposal by a proof derived from relativity theory, but we will not present the details of this proof.

De Broglie described his theoretical ideas in his doctoral thesis with the title "Recherches sur la Théorie des Quanta," which he presented in November 1924 to the Faculty of Natural Science at the Sorbonne. The examination committee of the faculty had difficulty believing the validity of de Broglie's proposals, and one of its members asked how they could be verified experimentally. De Broglie answered that such verification might be obtained by measuring the pattern of diffraction of a beam of electrons by a single crystal. The committee was unaware of the fact that such experiments had already been performed. They nevertheless awarded the doctor's degree to Louis de Broglie because they were impressed by the originality of his ideas.

De Broglie was clearly not a member of the inner circle of prominent theoretical physicists centered at Munich, Göttingen, Berlin, and Copenhagen. However, Einstein was made aware of de Broglie's work. Upon his request, de Broglie sent Einstein a copy of his thesis, which the latter read in December 1924. Einstein brought the thesis to the attention of Max Born in Göttingen, who in turn described it to his colleague in experimental physics *James Franck* (1882–1964). Franck remembered that a few years earlier, two scientists at the AT&T Research Laboratories, *Clinton Joseph Davisson* (1881–1958) and *Charles Henry Kunsman* (1890–1970), had measured the scattering of a beam of electrons by a platinum plate. The experimental results exhibited some features that could be interpreted as a diffraction pattern. One of Born's graduate students, *Walter Maurice Elsasser* (1904–1991), calculated the diffraction pattern due to interference of de Broglie

waves associated with the electron beam and found that his theoretical result agreed exactly with the experiments. These results and subsequent diffraction measurements provided convincing experimental evidence for the validity of de Broglie's theories.

De Broglie received the 1929 Nobel Prize in physics, and we quote the words of the chairman of the Nobel Committee for Physics:

When quite young you threw yourself into the controversy ranging round the most profound problem in physics. You had the boldness to assert, without the support of any known fact, that matter had not only a corpuscular nature, but also a wave nature. Experiment came later and established the correctness of your view. You have covered in fresh glory a name already crowned for centuries with honor.

De Broglie's theoretical model was primarily intended to describe the motion of free particles, but he also presented an application to the motion of an electron around a nucleus. In the latter case, it seems logical to assume that the circumference of the closed orbit should be equal to an integral number  $n$  of de Broglie wavelength  $h/p$ . This requirement is identical to Sommerfeld's quantization rule mentioned in Section 1.III.

About a year after the publication of de Broglie's thesis, Erwin Schrödinger formulated the definitive mathematical foundation of quantum mechanics in a series of six papers that he wrote in less than a year during 1926. The key feature of this model, the Schrödinger equation, has formed the basis for all atomic and molecular structure calculations ever since it was first proposed, and it is without doubt the best-known equation in physics.

There are actually two Schrödinger equations, the time-independent equation

$$-\frac{\hbar^2}{2m}\Delta\psi + V\psi = E\psi \quad (1-16a)$$

and the time-dependent one

$$-\frac{\hbar^2}{2m}\Delta\psi + V\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t} \quad (1-16b)$$

The two equations are closely related.

The symbol  $\Delta$  denotes the Laplace operator:

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (1-17)$$

and the symbol  $V$  represents the potential energy. The above are one-particle Schrödinger equations, but their generalization to larger systems containing more particles is quite straightforward.

We shall see that for positive values of energy  $E$ , the Schrödinger equation describes the motion of a free or quasi-free particle. In the case of bound particles

whose motion is confined to closed orbits of finite magnitude, additional restraints must be imposed on the solutions of the equations. Mathematicians classify the Schrödinger equation as a partial differential equation with boundary conditions that is similar in nature to the problem of a vibrating string. Schrödinger had received an excellent education in mathematics, and he was familiar with this particular topic.

Erwin Schrödinger was the most interesting of the physicists listed in Table 1-1, since he had both a fascinating personality and an interesting life. He was born in 1887 into a comfortable upper-middle-class family in Vienna, where his father owned a profitable business. He had a happy childhood. He was the top student at the most prestigious type of high school, the gymnasium, that he attended. He also took full advantage of the lively cultural atmosphere in Vienna at the time; he particularly liked the theatre. At school he was interested in literature and philosophy, but he decided to study theoretical physics and enrolled at the University of Vienna in 1906.

Erwin's academic studies in physics proceeded smoothly. He received an excellent education in mathematical physics and was awarded a doctor's degree in 1910. In order to satisfy his military obligations, he volunteered for officer's training and served for 1 year (1910–1911), after which he joined the army reserve. He returned to the university in 1911, and in 1914 he was admitted to the faculty of the University of Vienna as a *Privatdozent*. This meant that he was allowed to conduct research and give lectures at the university, but he did not necessarily receive any salary in return.

Schrödinger's academic career, like that of Louis de Broglie, was rudely interrupted by the outbreak of the First World War in 1914. Schrödinger was recalled to military duty and served as an artillery officer at the southern front until 1917. At that time he was reassigned as a meteorology officer in the vicinity of Vienna since he had taken a course in meteorology as a student. This transfer may very well have saved his life.

After the war Schrödinger returned home to Vienna, where he found that living conditions were quite bleak. This should not have come as a surprise; after all, Austria had lost the war. His father's business had failed as a result of the war and his savings had been eroded due to inflation, so the financial conditions of the family were far from favorable. His health also deteriorated, and he died towards the end of 1919.

During that time Erwin received a small stipend from the university, and even though this was inadequate to meet his living expenses, he worked very hard at research. His main interest was the theory of color, a subject that straddled physics, physiology, and psychology. He wrote a number of research papers on the subject followed by a highly regarded review.

Schrödinger's academic career took a turn for the better in 1920, when he was offered a low-level faculty position at the University of Jena. He did not stay there very long because, like most other professors, he was interested in finding a better job at a more prestigious university. In the next few years he moved first to Stuttgart, then to Breslau, and finally, in 1921, to Zürich, where he was appointed

a full professor of theoretical physics at a generous annual salary of 14,000 SFr. He had found an excellent academic position away from the political and economic turmoil of Germany and Austria.

At the time of his appointment, Schrödinger could be considered a typical average physics professor. He was highly knowledgeable and he had published a number of competent research papers, but his name was not associated with any major discovery. He was probably best known for his work on the theory of color.

In addition to the University of Zürich, there is also a technical university in the city, the Eidgenössische Technische Hochschule (E.T.H.), which was regarded as the more prestigious of the two. The physics professor at the E.T.H. was the Dutchman Pieter Debye, who was better known and more highly regarded than Schrödinger at the time. The two physics professors met frequently at their joint physics seminar.

It was Debye who first became aware of de Broglie's work and who brought it to Schrödinger's attention. De Broglie's thesis had been published in a French physics journal, and Debye suggested that Schrödinger give a seminar in order to explain it to the Zürich physicists. After the seminar Schrödinger realized that in classical physics waves were usually interpreted as solutions of a partial differential equation, the so-called wave equation. It occurred to him that it might be possible to formulate a similar wave equation for the description of de Broglie waves, and he set out to try to derive such a wave equation.

At first Schrödinger tried to derive a wave equation by using relativity theory. Even though he was successful, the results of this equation did not agree with the experimental information on the hydrogen atom's spectrum. This lack of agreement could probably be attributed to effects of the electron spin, which had not yet been discovered. Schrödinger nevertheless changed his approach, and he derived a wave equation based on classical nonrelativistic mechanics.

During the next 6 months, the first half of the year 1926, Schrödinger wrote six research papers in which he presented the complete mathematical foundation of nonrelativistic quantum mechanics. In fact, the contents of this book are based almost entirely on the six Schrödinger papers.

Schrödinger's wave equation bears some resemblance to the equation describing the motion of a vibrating string. The mathematicians classify it as a partial differential equation with boundary conditions. The solutions of the differential equation are required to assume specific values at various points. This condition is satisfied only for a discrete set of values of a parameter in the equation. The German word for these values is *Eigenwerte*, which has been translated in English as eigenvalues rather than the more suitable term specific values. This particular problem is also known among mathematicians as the Sturm-Liouville problem.

In the Schrödinger equation, the adjustable parameter is the energy and its eigenvalues correspond to the quantized stationary states for the system. In addition to proposing the equation, Schrödinger derived its solution for a variety of systems, including the hydrogen atom. He accomplished this during a 6-month period of intense concentration, a truly spectacular effort.

Schrödinger remained in Zürich until 1927, when he received an offer to become Planck's successor at the University of Berlin. The offer was hard to resist because the position was not only very lucrative but also extremely prestigious; the second chair of theoretical physics at the university was held by Einstein. Also, Berlin was a vibrant and attractive city at the time. Schrödinger had won what was probably the best academic job in Europe just before his fortieth birthday. All went well until 1933, when the Nazis under the leadership of Adolf Hitler came to power and introduced a succession of anti-Jewish laws. Einstein happened to be in the United States at the beginning of 1933, and he decided not to return to Germany. Schrödinger had never been particularly interested in politics, but in this instance he decided that he no longer wanted to stay in Germany. He moved to Oxford, where he became a Fellow of Magdalen College. He did not formally resign his professorship, but he requested a leave of absence and sent a postcard to the physics department to inform the students that his lectures for the fall semesters would be canceled. He did not make a dramatic exit; he just left.

During the next few years, Schrödinger traveled widely. He received the 1933 Nobel Prize in physics, and he was in great demand as a guest lecturer. He also had to find a permanent academic position since his appointment in Oxford was temporary. He had the choice of a number of academic positions, but he made an almost fatal error in accepting a professorship at the University of Graz in his native Austria. When the Austrian Nazis managed to arrange a merger with Germany, the so-called *Anschluss*, Schrödinger found himself suddenly in a very precarious position since his departure from Germany had deeply offended the Nazis. He was fortunate to be able to leave the country without being arrested, but he had to leave all his possessions, including his money and valuables, behind.

The president of Ireland invited Schrödinger to become the director of a newly established institute for theoretical physics in Dublin, where he spent the next 18 years. In 1956, when Schrödinger's health was already beginning to fail, he moved back to his native Vienna as a professor of physics. He died there in 1961.

## VII. THE FINAL TOUCHES OF QUANTUM MECHANICS

In Schrödinger's work the emphasis was on the energy eigenvalues, the discrete values of the energy parameter that correspond to acceptable solutions of the differential equation. It was shown that these eigenvalues coincide with the energies of Bohr's stationary states. Much less attention was paid to the physical interpretation of solutions of the equation corresponding to each eigenvalue; these latter functions became known as eigenfunctions.

It was Born who proposed in the same year, 1926, that the product of an eigenfunction  $\psi$  and its complex conjugate  $\psi^*$  represents the probability density of the particle. In other words, the probability of finding the particle in a small-volume element surrounding a given point is given by the product of the volume element and the value of the probability density  $\psi \cdot \psi^*$  at that point.

Born's interpretation is easily extended to situations where a one-particle system is described by a wave function  $\psi(x, y, z, t)$  that may or may not be an eigenfunction corresponding to a stationary state. Here the product  $\psi \cdot \psi^*$  is again a representation of the probability density of the particle. Even though quantum mechanics does not offer an exact prediction of the position of the particle, it offers an exact prediction of the statistical probability distribution of locating the particle.

Born first proposed the probabilistic interpretation of the wave function in relation to a theory of electron scattering, in particular the scattering of a high-energy electron by an atom. He later extended the idea to all other aspects of quantum mechanics, and it became universally accepted. Born's statistical interpretation of the wave function is probably the most important of his many contributions to the development of quantum mechanics. The award of the 1954 Nobel Prize in physics to Max Born at age 72 was motivated primarily by this contribution.

The formal description of quantum mechanics as we know it today was completed in just a few years. We briefly describe the various developments. The motion of an electron around a nucleus in an atom has often been compared to the motion of the planets around the sun. We know that the Earth not only describes an annual orbit around the sun but also performs a diurnal rotation around its axis. The idea occurred to two graduate students at Leiden University, *George Uhlenbeck* and *Samuel Goudsmit*, that by analogy, the electron might also be capable of rotational motion around its axis. At the time, there were certain features in atomic spectra (referred to as the anomalous Zeeman effect) that defied all logical explanation. Goudsmit and Uhlenbeck proposed in 1925 that the assumption of rotational motion within the electron and subsequent quantization of this motion offered the possibility of explaining the anomalous Zeeman effect. The rotational motion of the electron became known as the electron spin. Goudsmit and Uhlenbeck's theory led to perfect agreement with the experimental atomic spectral features.

Initially Goudsmit and Uhlenbeck's ideas were severely criticized because they appeared to be inconsistent with classical electromagnetic theory. However, early in 1926, *Lewellyn Hilleth Thomas* (1903–1992) showed that Goudsmit and Uhlenbeck's assumptions were entirely correct if the relativistic effect was taken into account.

The theoretical description of the spinning electron became a fundamental aspect of quantum mechanics when Paul Dirac generalized the Schrödinger equation to make it consistent with relativity theory. The existence of the electron spin was an essential feature of the Dirac equation. We should add that relativistic quantum mechanics is not included in this book since we believe it to be too sophisticated for our level of presentation.

The Schrödinger equation is easily extended to many-electron systems, but in that case the wave function is subject to an additional restraint due to the Pauli exclusion principle. In interpreting the electronic structure of an atom, it had been customary to assign each electron for identification purposes to a stationary state determined by a set of quantum numbers. In order to be consistent with the experimental information on atomic structure, Wolfgang Pauli imposed in 1925 the condition that no more than two electrons could be assigned to the same stationary

state. When the spin is included in the definition of the stationary state, no more than one electron can be assigned to each state. This condition became known as the Pauli exclusion principle. We will present a more general and more exact formulation of the exclusion principle when we discuss the helium atom in Chapter 10.

The mathematical formalism of quantum mechanics was completed in 1927, and all that remained was to find solutions to the Schrödinger equation for atomic and molecular systems. This required the introduction of approximate techniques since exact analytical solutions could be derived only for a limited number of one-particle systems. Today, highly accurate solutions of the Schrödinger equation for relatively large molecules can be obtained. This is due to the concerted effort of many scientists and also to the introduction of high-speed computers. We may conclude that the majority of the problems involving the application of quantum mechanics to atomic and molecular structure calculations have been solved.

## VIII. CONCLUDING REMARKS

Quantum mechanics is basically a conglomerate of revolutionary new ideas and concepts. The most important of these are Planck's quantization, Bohr's introduction of stationary states, Heisenberg's uncertainty relations, de Broglie's wave-particle duality, Schrödinger's equation, and Born's statistical interpretation of the wave function. Dirac remarked in his textbook that these new theoretical ideas are built up from physical concepts that cannot be explained in terms of things previously known to the student and that cannot be explained adequately in words at all. They definitely cannot be proved.

It is best to look upon them as new fundamental laws of physics that form a logically consistent structure and that are necessary to interpret all known experimental facts.

We have made a deliberate attempt to present these novel ideas from a non-mathematical perspective. Unfortunately, it is not possible to apply the ideas without making use of mathematical techniques. We present the necessary background material in the following two chapters.



---

# 2

---

## THE MATHEMATICS OF QUANTUM MECHANICS

### I. INTRODUCTION

We have seen that in previous times a successful theoretical physicist had to have a broad and thorough knowledge of mathematics. For many physicists mathematics was nothing more than a useful tool, but some physicists made fundamental contributions to mathematics; for instance, Isaac Newton created calculus. David Hilbert, who was one of the most prominent and creative mathematicians of the nineteenth and twentieth centuries, became interested in physics later in life. We mentioned that Hilbert and his student, Richard Courant, published in 1924 a mathematics textbook called *Methods of Mathematical Physics*, which by a happy coincidence contained many of the mathematical topics that were necessary for understanding quantum mechanics.

All of the scientists listed in Table 1-1 were expert mathematicians, but one of them, Arnold Sommerfeld, was actually a respected professor of mathematics who had made important original contributions to the field before he became interested in physics. It may be interesting to give a brief description of his career.

Sommerfeld was born and raised in Königsberg, the capital of what was then called East Prussia. He studied mathematics at the University of Königsberg and graduated with a doctor's degree in 1891. During his studies he attended a number of lectures by David Hilbert, who was also born in Königsberg and had just been appointed a *Privatdozent* (assistant professor) at the university. It is worth noting that Hilbert and Sommerfeld became and remained close friends. After graduating,

Sommerfeld continued his mathematics studies with *Felix Klein* (1849–1925) in Göttingen. In addition to being an outstanding mathematician, Klein was a first-rate administrator and politician, and at the time he was probably the most influential mathematician in the country. Sommerfeld became Klein's star student and he tackled one of the more challenging problems in mathematical physics, the motion of the gyroscope. Sommerfeld's elegant solution of the problem, which was published between 1897 and 1910, was considered a major contribution to the field of mathematics.

In 1900 Sommerfeld became a professor of mechanics at the Technical University of Aachen, where he became interested in practical applications of mathematics. This led in 1906 to his appointment to the chair of theoretical physics at the University of Munich, where he remained until his death in 1951 as a result of an automobile accident.

Sommerfeld made some important contributions to the development of quantum mechanics. He also turned out to be an outstanding and popular teacher. Three of his students—Heisenberg, Debye, and Pauli—were Nobel Prize recipients, and many others became prominent physicists. Sommerfeld himself never received a Nobel Prize even though he was nominated numerous times. It recently became known that this may be attributed to the opposition of *Carl Wilhelm Oseen* (1879–1944), who was for many years the chairman of the Nobel Prize physics committee.

Two important mathematical disciplines that are essential for understanding quantum mechanics are linear algebra (matrices and determinants) and differential equations. We present both of these topics in this chapter. We will show that all differential equations discussed in this book may be derived from only two particular types, and we limit our discussion to those two equations.

Other mathematical topics will be presented throughout the book, where they will be linked to the corresponding features and applications of quantum mechanics. For instance, Fourier analysis and the mathematical description of waves will be discussed in Chapter 5 in combination with the wave mechanics of the free particle. In this way, the relevance of these various mathematical topics is better illustrated.

## II. DIFFERENTIAL EQUATIONS

The majority of differential equations encountered in theoretical physics are linear second-order differential equations of the type

$$\frac{d^2u}{dx^2} + p(x)\frac{du}{dx} + q(x)u = 0 \quad (2-1)$$

This equation always has two linearly independent solutions,  $u_1(x)$  and  $u_2(x)$ , and its general solution may be represented as

$$u(x) = Au_1(x) + Bu_2(x) \quad (2-2)$$

A standard technique for solving the differential equation (2-1) consists of substituting a power series expansion for the functions  $u(x)$ . In some cases the coefficients of the power series may then be derived in a straightforward manner, while in other cases this technique may not be effective. A second approach for dealing with Eq. (2-2) involves its transformation into one of the standard differential equations whose solutions have been extensively studied. Those solutions are known as special functions. They are usually named after the mathematicians who first studied them, and they are tabulated and described in detail in the mathematical literature.

We are fortunate that all solutions of the Schrödinger equation that we discuss in this book are either of a trivial nature or may be reduced to just one type of special function. The latter function was first introduced in 1836 by the German mathematician *Ernst Edward Kummer* (1810–1893) and is known as the confluent hypergeometric function  ${}_1F_1(a, b; x)$  or as Kummer's function. It is defined as the power series

$${}_1F_1(a, b; x) = 1 + \frac{a x}{b 1!} + \frac{a(a + 1)x^2}{b(b + 1) 2!} + \frac{a(a + 1)(a + 2)x^3}{b(b + 1)(b + 2) 3!} + \dots, \text{ etc.} \quad (2-3)$$

We discuss this function in the following section.

The equations of a trivial nature that we alluded to are those in which the functions  $p(x)$  and  $q(x)$  are constants. In the latter case the solutions are obtained by substituting

$$u(x) = \exp(\mu x) \quad (2-4)$$

into Eq. (2-1), which leads to the quadratic equation

$$\mu^2 + p\mu + q = 0 \quad (2-5)$$

This equation has two roots,  $\mu_1$  and  $\mu_2$ , and the general solution of the equation is obtained as

$$u(x) = A \exp(\mu_1 x) + B \exp(\mu_2 x) \quad (2-6)$$

In the special case where the two roots coincide, the solution becomes

$$u(x) = (A + Bx) \exp(\mu_1 x) \quad (2-7)$$

### III. KUMMER'S FUNCTION

In order to illustrate the series expansion method for the solution of second-order differential equations, we apply it to the differential equation corresponding to

Kummer's function. It has the form

$$x \frac{d^2 u}{dx^2} + (b - x) \frac{du}{dx} - au = 0 \quad (2-8)$$

This differential equation is solved by substituting the following power series expansion:

$$u(x) = x^\rho \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} c_n x^{n+\rho} \quad (2-9)$$

Substitution gives

$$\sum_{n=0}^{\infty} (n + \rho)(n + \rho + b - 1) c_n x^{n+\rho-1} - \sum_{n=0}^{\infty} (n + \rho + a) c_n x^{n+\rho} = 0 \quad (2-10)$$

or

$$\rho(\rho + b - 1) c_0 x^{\rho-1} + \sum_{n=0}^{\infty} [(n + \rho + 1)(n + \rho + b) c_{n+1} - (n + \rho + a) c_n] x^{n+\rho} = 0 \quad (2-11)$$

We now impose the condition that each coefficient of the power series is equal to zero; in that case, the power series is a solution of the differential equation. The condition for the first term of the series, corresponding to  $x^{\rho-1}$ , is

$$\rho(\rho + b - 1) = 0 \quad (2-12)$$

This equation has two solutions

$$\rho_1 = 0 \quad \rho_2 = 1 - b \quad (2-13)$$

corresponding to the two linearly independent solutions  $u_1$  and  $u_2$  of the equation.

We first consider the solution  $u_1(x)$ , which we obtain from Eq. (2-11) by substituting  $\rho = 0$ . We find

$$c_{n+1} = \frac{(n + a)}{(n + b)(n + 1)} c_n \quad (2-14)$$

It follows that the series expansion method is convenient in the present case since it enables us to derive the values of the coefficients sequentially. We may set  $c_0$  equal to unity, and we then find

$$c_0 = 1 \quad c_1 = \frac{a}{b} \frac{1}{1!} \quad c_2 = \frac{a(a+1)}{b(b+1)} \frac{1}{2!} \quad c_3 = \frac{a(a+1)(a+2)}{b(b+1)(b+2)} \frac{1}{3!} \dots, \text{etc.} \quad (2-15)$$

This result is identical to the definition (2-3) of Kummer's function, and we find therefore that

$$u_1(x) = {}_1F_1(a, b; x) \tag{2-16}$$

The second solution of the differential equation may also be expressed in terms of the Kummer function. By substituting  $\rho_2$  into Eq. (2-11) we obtain

$$u_2(x) = x^{1-b} {}_1F_1(a - b + 1, 2 - b; x) \tag{2-17}$$

and

$$u(x) = A {}_1F_1(a, b; x) + B x^{1-b} {}_1F_1(a - b + 1, 2 - b; x) \tag{2-18}$$

where  $A$  and  $B$  are two arbitrary undetermined parameters.

The confluent hypergeometric function has been studied extensively, but we mention only one of its main properties. By substituting

$$u(x) = e^x w(x) \tag{2-19}$$

into the differential equation (2-8), it is possible to derive Kummer's relation

$${}_1F_1(a, b; x) = e^x {}_1F_1(b - a, b; -x) \tag{2-20}$$

Asymptotic expansions for Kummer's function have also been derived, and we mention the result that for large values of the variable  $x$  the function behaves asymptotically as the exponential function  $e^x$ . A different asymptotic behavior is found when the parameter  $a$  happens to be a negative integer because in that case the function is reduced to a finite polynomial of the variable  $x$ .

The above brief survey covers all the aspects of the theory of differential equations that are needed in this book. In the rest of this chapter we will discuss some relevant features of linear algebra, the theory of matrices and determinants.

#### IV. MATRICES

A matrix is defined as a two-dimensional rectangular array of numbers (or functions) that are called the elements of the matrix. A matrix may be represented as follows:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \dots \dots & a_{1,N} \\ a_{2,1} & a_{2,2} & a_{2,3} \dots \dots & a_{2,N} \\ a_{3,1} & a_{3,2} & a_{3,3} \dots \dots & a_{3,N} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{M,1} & a_{M,2} & a_{M,3} \dots \dots & a_{M,N} \end{bmatrix} \tag{2-21}$$

An abbreviated notation for a matrix is

$$\begin{aligned} [a_{i,j}] \quad & i = 1, 2 \dots M \\ & j = 1, 2 \dots N \end{aligned} \quad (2-22)$$

or simply

$$[\mathbf{A}] \quad \text{or} \quad \mathbf{A} \quad (2-23)$$

We note that a vector is represented by a boldface lowercase letter, whereas a matrix is described by a boldface capital letter in square brackets or by a boldface capital letter alone. The notation (2-23) is allowed only after the matrix has been defined in more detail.

All elements of a matrix that are on the same horizontal level are said to form a row, and all elements that are on the same vertical line are said to form a column.

The product  $[\mathbf{C}]$  of two matrices  $[\mathbf{A}]$  and  $[\mathbf{B}]$  is obtained by multiplying the rows of the first matrix  $[\mathbf{A}]$  with columns of the second matrix  $[\mathbf{B}]$ ; in other words

$$c_{i,k} = \sum_j a_{i,j} b_{j,k} \quad (2-24)$$

The multiplication is feasible only if the number of elements in the rows of the first matrix is equal to the number of elements in the columns of the second matrix or, in other words if the length of the first matrix is equal to the height of the second matrix.

It may also be seen that the product  $[\mathbf{A}] \cdot [\mathbf{B}]$  is not necessarily the same as the product  $[\mathbf{B}] \cdot [\mathbf{A}]$ . In the special case where the two products are equal, the two matrices are said to commute.

It may be helpful to present a simple application of the use of matrices, namely, their description of coordinate transformations. We consider an  $N$ -dimensional vector  $\mathbf{u}$  with components  $(u_1, u_2, \dots, u_N)$ . We now consider a coordinate transformation described by a square matrix  $[b_{j,k}]$  that transforms  $\mathbf{u}$  into  $\mathbf{v}$ . This may be represented by the following matrix multiplication:

$$\begin{bmatrix} b_{1,1} & b_{1,2} & \dots & \dots & b_{1,N} \\ b_{2,1} & b_{2,2} & \dots & \dots & b_{2,N} \\ \dots & \dots & \dots & \dots & \dots \\ b_{N,1} & b_{N,2} & \dots & \dots & b_{N,N} \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_N \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_N \end{bmatrix} \quad (2-25)$$

or as

$$v_j = \sum_{k=1}^N b_{j,k} u_k \quad (2-26)$$

We can now subject the vector  $\mathbf{v}$  to a second transformation represented by a square matrix  $[a_{i,j}]$ . The result is a vector  $\mathbf{w}$  given by

$$w_i = \sum_{j=1}^N a_{i,j} v_j = \sum_{j=1}^N \sum_{k=1}^N a_{i,j} b_{j,k} u_k \quad (2-27)$$

It follows that the transformation from  $\mathbf{u}$  to  $\mathbf{w}$  may also be represented by a square matrix  $[c_{i,k}]$  that is defined as

$$c_{i,k} = \sum_{j=1}^N a_{i,j} b_{j,k} \quad (2-28)$$

This is consistent with our previous definition of matrix multiplication, where  $[\mathbf{C}]$  is the product of  $[\mathbf{A}]$  and  $[\mathbf{B}]$

In quantum mechanics we deal almost exclusively with square matrices. We present a number of definitions and features of this class of matrices. We first introduce the Kronecker symbol  $\delta_{i,j}$ , which is defined as

$$\begin{aligned} \delta_{i,j} &= 1 & \text{if } i &= j \\ \delta_{i,j} &= 0 & \text{if } i &\neq j \end{aligned} \quad (2-29)$$

We define a diagonal matrix as a matrix having nonzero elements on its diagonal only. Its elements may be written as

$$a_{i,j} = a_i \delta_{i,j} \quad (2-30)$$

A special case is the unit matrix, denoted by the symbol  $\mathbf{I}$  and defined by

$$a_{i,j} = \delta_{i,j} \quad (2-31)$$

A symmetric matrix is defined by the condition

$$a_{i,j} = a_{j,i} \quad (2-32)$$

whereas a skew-symmetric matrix is defined by

$$a_{i,j} = -a_{j,i} \quad (2-33)$$

A very important type of matrix in quantum mechanics is the Hermitian matrix. We should realize that the elements of a matrix may in general be complex quantities. A Hermitian matrix is then defined by the condition

$$a_{i,j} = a_{j,i}^* \quad (2-34)$$

This differs from the definition of a symmetric matrix.

Finally, we mention unitary matrices  $U$ , which are defined by the conditions

$$\begin{aligned} \sum_k u_{i,k} u_{j,k}^* &= \delta_{i,j} \\ \sum_k u_{k,i}^* u_{k,j} &= \delta_{i,j} \end{aligned} \tag{2-35}$$

### V. PERMUTATIONS

The definition of a determinant is dependent upon the properties of permutations, so we discuss these first. In order to describe permutations, we consider a finite set of quantities that are identified by positive-integer numbers. We assume that these quantities are arranged in a monotonic increasing sequence of their numbers  $(1, 2, 3 \dots, N)$ . Scrambling the objects into a different sequence characterized by the set of numbers  $(n_1, n_2, n_3, \dots, n_N)$  is called a permutation of the original sequence.

Any permutation may be represented as the result of a sequence of pairwise interchanges or swaps. We present the following example, where at each stage we have underlined the two numbers that are swapped:

$$\begin{array}{cccccccc} \underline{9} & 6 & \underline{1} & 2 & 4 & 5 & 8 & 7 & 3 \\ 1 & \underline{6} & 9 & \underline{2} & 4 & 5 & 8 & 7 & 3 \\ 1 & 2 & \underline{9} & 6 & 4 & 5 & 8 & 7 & \underline{3} \\ 1 & 2 & 3 & \underline{6} & \underline{4} & 5 & 8 & 7 & 9 \\ 1 & 2 & 3 & 4 & \underline{6} & \underline{5} & 8 & 7 & 9 \\ 1 & 2 & 3 & 4 & 5 & 6 & \underline{8} & \underline{7} & 9 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \tag{2-36}$$

The above permutation  $P$  is the result of six swaps, an even number. It is customary to introduce a symbol  $\delta_p$ , which is defined as

$$\begin{aligned} \delta_p &= 1 && \text{if } P \text{ is even} \\ \delta_p &= -1 && \text{if } P \text{ is odd} \end{aligned} \tag{2-37}$$

The permutation is even if it is the result of an even number of swaps, and it is odd if it is the result of an odd number of swaps.

It is also useful to know the total number  $S(N)$  of possible permutations of  $N$  objects or the number of ways that  $N$  objects may be arranged. This number is given by

$$S(N) = 1 \cdot 2 \cdot 3 \cdot 4 \dots N = N! \tag{2-38}$$



It is easily seen that in a given sequence of  $N$  objects, an additional object may be inserted or added in  $(N + 1)$  different positions so that

$$S(N + 1) = (N + 1)S(N) \quad (2-39)$$

which immediately leads to the expression (2-38).

## VI. DETERMINANTS

A determinant is an entity derived from the elements of a square matrix by following a given well-defined procedure. We denote the elements of the square matrix of order  $N$  by the symbols  $a(i, j)$ . In order to derive the determinant, we construct all possible products of the elements  $a(i, j)$  that are subject to the condition that no two elements in each product belong to the same column or the same row. We may write these products in two different ways, namely, as

$$a(1, n_1) \cdot a(2, n_2) \cdot a(3, n_3) \dots a(N, n_N) \quad (2-40)$$

or as

$$a(n_1, 1) \cdot a(n_2, 2) \cdot a(n_3, 3) \dots a(n_N, N) \quad (2-41)$$

with the understanding that all the numbers  $n_i$  are different from each other.

Each permutation of the set of indices  $(n_1, n_2, n_3, \dots, N)$  produces a different product, and since there is a total of  $N!$  permutations according to Eq. (2-38), there is also a total of  $N!$  different products.

The determinant  $\Delta$  is now defined in two alternative equivalent ways as the sum of all possible products (2-40) or (2-41), where each term is multiplied by a factor  $\delta_P$ . The latter coefficient is defined according to Eq. (2-37) as plus unity when the permutation  $P$  of the indices is even and as minus unity when the permutation is odd. We may represent these two equivalent definitions of the determinant by the following mathematical expressions:

$$\Delta = \sum_P \delta_P P(n_1, n_2, \dots, n_N) [a(1, n_1) \cdot a(2, n_2) \dots a(N, n_N)] \quad (2-42)$$

and

$$\Delta = \sum_P \delta_P P(n_1, n_2, \dots, n_N) [a(n_1, 1) \cdot a(n_2, 2) \dots a(n_N, N)] \quad (2-43)$$

It is easily seen that both expressions (2-42) and (2-43) are equal to one another since they contain exactly the same terms except that the latter are listed in a different order.

There is some controversy about the discovery of determinants, but it is now believed that they were first mentioned in 1693 in a letter by the German mathematician *Gottfried Wilhelm Leibnitz* (1646–1716).

A determinant is often represented in the literature as its matrix with vertical lines instead of square brackets:

$$A = \begin{vmatrix} a_{1,1} & a_{1,2} \dots & \dots & \dots & \dots & a_{1,N} \\ a_{2,1} & a_{2,2} \dots & \dots & \dots & \dots & a_{2,N} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{N,1} & a_{N,2} \dots & \dots & \dots & \dots & a_{N,N} \end{vmatrix} \quad (2-44)$$

As illustrations we list the values of determinants of order 2 by 2

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 \quad (2-45)$$

and of order 3 by 3

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - a_3 b_2 c_1 - b_3 c_2 a_1 - c_3 a_2 b_1 \quad (2-46)$$

The evaluation of larger determinants may become very laborious; for example, a determinant of order 6 by 6 contains 720 products. On the other hand, such calculations are easily programmed, and they are conveniently done with the aid of electronic computers. We should also realize that the great advantage of using determinants in theoretical derivations is due to the fact that in most cases there is no need to evaluate each determinant.

## VII. PROPERTIES OF DETERMINANTS

It follows from the definitions (2-42) and (2-43) of a determinant that we obtain the same terms if we switch either a pair of columns or a pair of rows in a determinant. However, each term then has the opposite sign because of the changes in the factors  $\delta_p$ . We conclude that an exchange of a pair of columns or a pair of rows leads to a change of sign in the determinant. It also follows that a determinant with two identical rows or two identical columns is equal to zero.

Multiplication of a column or row of a determinant by a constant  $\lambda$  results in the multiplication of the determinant by  $\lambda$ . Consequently, a determinant is also zero if a pair of rows or columns are proportional to each other. It may be shown in general that a determinant is equal to zero if there is a linear relationship between either its rows or its columns.

An effective procedure for evaluating a determinant is by means of its expansion. Let us, for instance, consider the determinant  $A$  of Eq. (2-44), and let us identify all contributions containing the element  $a_{1,1}$ . The sum of these terms may be represented as a determinant that may be written as

$$A_{1,1} = \begin{vmatrix} a_{2,2} & a_{2,3} & \dots & a_{2,N} \\ a_{3,2} & a_{3,3} & \dots & a_{3,N} \\ \dots & \dots & \dots & \dots \\ a_{N,2} & a_{N,3} & \dots & a_{N,N} \end{vmatrix} \quad (2-47)$$

The determinant  $A_{1,1}$  is derived from  $A$  by deleting both the column and the row containing the element  $a_{1,1}$  and it is called the minor of the element  $a_{1,1}$ .

In a similar way, we may identify all products that contain the element  $a_{1,2}$ . We may accomplish this by switching the first and second columns of the determinant. The sum of all terms of  $A$  containing the element  $a_{1,2}$  is then the product of  $a_{1,2}$  and its minor,  $A_{1,2}$ . The minor  $A_{1,2}$  is obtained by deleting the row and column containing  $a_{1,2}$  from  $A$  and multiplying by  $(-1)$  since the switch of columns 1 and 2 changed the sign of the coefficients  $\delta_p$ .

In general, we may now define the set of minors  $A_{1,k}$  of the first row of the determinant  $A$  by deleting in each case the first row and the  $k$ th column from  $A$  and multiplying by  $(-1)^{k+1}$ :

$$A_{1,k} = (-1)^{k+1} \cdot \begin{vmatrix} a_{1,1} \dots \dots & \dots & a_{1,k-1} & a_{1,k+1} \dots \dots & a_{1,N} \\ a_{2,1} \dots \dots & \dots & a_{2,k-1} & a_{2,k+1} \dots \dots & a_{2,N} \\ \dots & \dots & \dots & \dots & \dots \\ a_{N,1} \dots \dots & \dots & a_{N,k-1} & a_{N,k+1} \dots \dots & a_{2,N} \end{vmatrix} \quad (2-48)$$

It can now easily be seen that the value  $\Delta$  of the determinant  $A$  of Eq. (2-44) is given by

$$\Delta = \sum_{k=1}^N a_{1,k} A_{1,k} \quad (2-49)$$

We may repeat this expansion for a different determinant where the top row has been replaced by one of its other rows. The result is

$$\sum_{k=1}^N a_{i,k} A_{1,k} = 0 \quad (2-50)$$

The value of the modified determinant is zero since two of its rows are identical. The above relation may be generalized to

$$\sum_{k=1}^N a_{i,k} A_{j,k} = \delta_{i,j} \cdot \Delta \quad (2-51)$$

The above relation will prove to be useful for deriving the solution of sets of linear equations.

Determinants may be evaluated by making use of the expansion (2-49) or by manipulating the rows or columns in order to create a row or column where all or most of the elements are zero. We illustrate the first approach by evaluating the determinant

$$D_5 = \begin{vmatrix} x & 1 & 0 & 0 & 0 \\ 1 & x & 1 & 0 & 0 \\ 0 & 1 & x & 1 & 0 \\ 0 & 0 & 1 & x & 1 \\ 0 & 0 & 0 & 1 & x \end{vmatrix} \quad (2-52)$$

We expand along the first row and obtain

$$D_5 = x \cdot \begin{vmatrix} 1 & x & 1 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & 1 & x \end{vmatrix} - \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & x & 1 & 0 \\ 0 & 1 & x & 1 \\ 0 & 0 & 1 & x \end{vmatrix} \quad (2-53)$$

Further expansion yields

$$D_5 = x^2 \cdot \begin{vmatrix} x & 1 & 0 \\ 1 & x & 1 \\ 0 & 1 & x \end{vmatrix} - x \begin{vmatrix} 1 & 1 & 0 \\ 0 & x & 1 \\ 0 & 1 & x \end{vmatrix} - \begin{vmatrix} x & 1 & 0 \\ 1 & x & 1 \\ 0 & 1 & x \end{vmatrix} \quad (2-54)$$

or

$$D_5 = (x^2 - 1)(x^3 - 2x) - x(x^2 - 1) = x(x^2 - 1)(x^2 - 3) \quad (2-55)$$

where we have made use of Eq. (2-46). It should be noted that the result of our evaluation is a function rather than a number. That is the reason we were careful to use the term *entity* in our definition of determinants at the beginning of Section 2.VI.

The following is an example of the second approach, namely, the evaluation of a determinant by adding rows and columns:

$$\begin{vmatrix} 1 & -1 & 4 & 0 & 2 \\ 0 & 1 & 4 & 2 & -1 \\ -1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 2 & -1 & 1 \\ 1 & -2 & 2 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 4 & 0 & 2 \\ 0 & 1 & 4 & 2 & -1 \\ 0 & 1 & 4 & 1 & 4 \\ 0 & 0 & 2 & -1 & 1 \\ 0 & -1 & -2 & 1 & -3 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 2 & -1 \\ 1 & 4 & 1 & 4 \\ 0 & 2 & -1 & 1 \\ -1 & -2 & 1 & -3 \end{vmatrix} \\ = \begin{vmatrix} 1 & 4 & 2 & -1 \\ 0 & 0 & -1 & 5 \\ 0 & 2 & -1 & 1 \\ 0 & 2 & 3 & -4 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 5 \\ 2 & -1 & -1 \\ 2 & 3 & -4 \end{vmatrix} = 34 \quad (2-56)$$

**VIII. LINEAR EQUATIONS AND EIGENVALUES**

The study of linear equations always begins by considering the case of  $N$  homogeneous equations with  $N$  variables. This is the simplest of the many different types of linear equations, and it is also a suitable starting point for deriving the solutions of all the other equations.

We consider the following set of linear homogeneous equations:

$$\begin{aligned}
 a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \dots \dots + a_{1,N}x_N &= 0 \\
 a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + \dots \dots + a_{2,N}x_N &= 0 \\
 a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 + \dots \dots + a_{3,N}x_N &= 0 \\
 \dots \quad \dots \quad \dots \quad \dots &= 0 \\
 \dots \quad \dots \quad \dots \quad \dots &= 0 \\
 a_{N,1}x_1 + a_{N,2}x_2 + a_{N,3}x_3 + \dots \dots + a_{N,N}x_N &= 0
 \end{aligned}
 \tag{2-57}$$

These equations have an obvious solution where every unknown  $x_i$  is equal to zero, but this so-called zero solution is always discounted. It should also be noted that any solution  $(x_1, x_2, \dots x_N)$  of the equation may be multiplied by an arbitrary parameter. A solution of the set of equations consists therefore of a ratio between the  $N$  variables rather than their absolute values. In general, the set of equations (2-57) will not have a solution since there are  $N$  equations and only  $(N - 1)$  independent variables.

We derive the solution of Eq. (2-57) by observing that the coefficients  $a_{i,j}$  define a determinant. We define the value of this determinant as  $\Delta$ , and we also define the minors  $A_{i,j}$  of the elements  $a_{i,j}$  of the determinant. According to Eqs. (2-49) and (2-50), we then have

$$\begin{aligned}
 \sum_k a_{1,k} A_{1,k} &= \Delta \\
 \sum_k a_{i,k} A_{1,k} &= 0 \quad i \neq 1
 \end{aligned}
 \tag{2-58}$$

It follows that

$$x_1 = A_{1,1} \quad x_2 = A_{1,2} \dots \dots \quad x_N = A_{1,N}
 \tag{2-59}$$

is a solution of the set of equations (2-57) if

$$\Delta = 0
 \tag{2-60}$$

If the determinant  $\Delta$  is different from zero, then the equations do not have a solution other than the zero solution.

Various other types of linear equations, such as  $N$  homogeneous equations with  $(N + 1)$  variables, or inhomogeneous equations may all be reduced to the set of equations (2-57).

In order to proceed to eigenvalue problems, we consider the following set of equations:

$$\begin{aligned}
 (a_{1,1} - \lambda)x_1 + a_{1,2}x_2 + a_{1,3}x_3 + \dots + a_{1,N}x_N &= 0 \\
 a_{2,1}x_1 + (a_{2,2} - \lambda)x_2 + a_{2,3}x_3 + \dots + a_{2,N}x_N &= 0 \\
 a_{3,1}x_1 + a_{3,2}x_2 + (a_{3,3} - \lambda)x_3 + \dots + a_{3,N}x_N &= 0 \\
 \dots \quad \dots \quad \dots \quad \dots & \dots = 0 \\
 a_{N,1}x_1 + a_{N,2}x_2 + a_{N,3}x_3 + \dots + (a_{N,N} - \lambda)x_N &= 0
 \end{aligned}
 \tag{2-61}$$

This is a set of  $N$  homogeneous equations with  $N$  unknowns, and according to our previous discussion, this set of equations will have a solution if and only if the determinant of the coefficients  $\Delta$  is equal to zero.

Let us now proceed to the eigenvalue problem of the matrix  $[a_{i,j}]$ . The problem is defined as the derivation of its eigenvalues and eigenvector. An eigenvalue  $\lambda_k$  is simply a value of the parameter  $\lambda$  for which the set of equations (2-61) has a non-zero solution and the corresponding nonzero solution  $x_1(k), x_2(k), \dots, x_N(k)$  is called the eigenvector associated with  $\lambda_k$ .

We have just concluded that the set of equations (2-61) has a nonzero solution only if the determinant of its coefficients is zero; in other words if

$$|a_{i,j} - \lambda\delta_{i,j}| = 0 \tag{2-62}$$

It is easily seen that this is a polynomial of order  $N$  in the parameter  $\lambda$ , and Eq. (2-62) therefore has  $N$  roots, which we denote by  $\lambda_1, \lambda_2, \dots, \lambda_N$ . The solution belonging to the root  $\lambda_k$  may be denoted by the vector  $\mathbf{x}(k)$ . The solutions of Eq. (2-61) may now be written in the form

$$[a_{i,j} - \lambda_k\delta_{i,j}]\mathbf{x}(k) = 0 \tag{2-63}$$

or as

$$[a_{i,j}] \cdot \mathbf{x}(k) = \lambda_k\mathbf{x}(k) \tag{2-64}$$

It is customary to call  $\lambda_k$  an eigenvalue of the square matrix  $[a_{i,j}]$  and  $\mathbf{x}(k)$  its corresponding eigenvector. If  $\mathbf{x}(k)$  is an eigenvector of the matrix  $[a_{i,j}]$ , then multiplication of  $[a_{i,j}]$  and  $\mathbf{x}(k)$  yields the same vector multiplied by a constant, the eigenvalue  $\lambda_k$ . It appears that we will see the same general pattern when we discuss the properties of operators in Chapter 5.

Nowadays the derivation of the eigenvalues and eigenvectors of large square matrices has become an integral part of computer programs for the calculation of

atomic and molecular structures. The mathematical process described in this section is not adopted in any of these computer programs since it is much too cumbersome and inefficient to be of any practical use. The only reason we presented the procedure is that it serves to define and explain the eigenvalue problem.

A great deal of effort has been invested in designing efficient programs for deriving the eigenvalues and eigenfunctions of large matrices. These efforts have been quite successful because it is now possible to derive the eigenvalues of very large matrices of orders of more than 1 million. Most of the programs contain procedures that have the goal of transforming the matrix to diagonal form since the elements of a diagonal matrix are equal to its eigenvalues. They are all very sophisticated and efficient computer programs, and they are so complex that we will not attempt to discuss them in detail.

**IX. PROBLEMS**

**2-1** Derive the general solution of the differential equation

$$\frac{d^2y}{dx^2} - 2\frac{dy}{dx} - 8y = 0$$

**2-2** Derive the general solution of the differential equation

$$\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = 0$$

**2-3** Prove that

$$\frac{d}{dx} {}_1F_1(a; c; x) = \frac{a}{c} {}_1F_1(a + 1; c + 1; x)$$

**2-4** Express the integral

$$\int_0^x e^t {}_1F_1(a; c; -t) dt$$

in terms of Kummer's function

**2-5** The function  ${}_1F_1(b + 2; b; x)$  is zero for two values of  $x$ , namely  $x_1$  and  $x_2$ . Determine these two values expressed in terms of the parameter  $b$ .

**2-6** Determine the values of  $\lambda$  for which the following homogeneous set of equations has a solution.

$$\begin{aligned} \lambda x + y + u &= 0 \\ \lambda y + z &= 0 \\ x - u &= 0 \\ y - 2z &= 0 \end{aligned}$$

- 2-7** Determine whether the following permutation of eight integer number is even or odd

$$(3, 5, 8, 6, 1, 7, 2, 4)$$

- 2-8** Determine the value of  $\lambda$  for which the following set of homogeneous equations has a solution and derive the solution

$$x + y + 2z = 0$$

$$x - y - z = 0$$

$$\lambda x - y + z = 0$$



---

# 3

---

## CLASSICAL MECHANICS

### I. INTRODUCTION

All theoretical predictions of classical mechanics are derived from just a few fundamental laws of nature. These laws are based on such compelling logic that nobody would dare to question their validity. They are also in agreement with all known experimental facts. Once a fundamental law of nature has been proposed and generally accepted, there is no further need to prove it.

The most widely known principle in all of the natural sciences is the fundamental law of conservation of energy. It states that in a closed system the sum of all types of energy is always a constant. This law has universal validity; it is applicable even to biological systems such as the human body. In the latter case, it is necessary to have a precise definition of the closed system that we consider and to include all possible types of energy that are involved in order to verify that the total energy of the system indeed remains constant.

The second fundamental law of classical mechanics may be attributed to the Italian mathematician *Gallileo Gallilei* (1564–1642) and to Isaac Newton. It states first that an object that is not subject to outside forces moves with a constant velocity  $v$ . If the object is subjected to a force  $F$ , the velocity will change and the acceleration  $a$  is proportional to the force:

$$F = m \cdot a \tag{3-1}$$

The proportionality constant is defined as the mass  $m$  of the object; it is sometimes referred to as the inert mass. We know from our own experience that the heavier an object is, the larger the force is that is required to produce acceleration.

Even though the basic rules of classical mechanics appear to be quite simple, their application in predicting the behavior of complex systems requires the use of more sophisticated mathematical disciplines. We assume that the reader is familiar with one of these disciplines, namely, calculus, which was formulated independently by Newton and by Leibnitz.

We have discussed some mathematical disciplines that are relevant to quantum mechanics in the previous chapter. Here we present some additional mathematical techniques that were developed as part of classical mechanics since they are also used in quantum mechanics.

## II. VECTORS AND VECTOR FIELDS

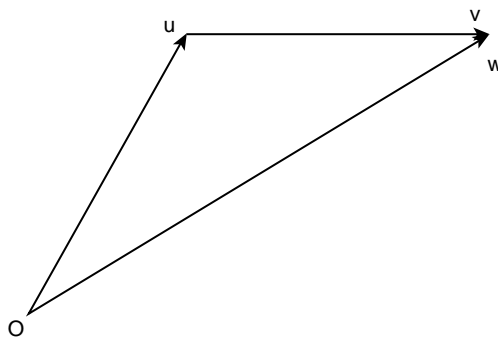
The first of these mathematical topics that we outline is vector analysis, that is, the study of vectors and vector fields, because they play an important role in both classical and quantum mechanics.

A vector  $\mathbf{v}$  is defined by mathematicians as a directed line segment, and it is denoted by a boldface symbol. Ordinary numbers are called scalars to emphasize their difference from vectors.

It is customary to define a vector  $\mathbf{v}$  by means of its three projections  $v_x$ ,  $v_y$ , and  $v_z$  on the X, Y, and Z axes of a Cartesian coordinate systems.

The sum  $\mathbf{w}$  of two vectors  $\mathbf{u}$  and  $\mathbf{v}$  is obtained by placing them end to end (see Figure 3-1). It is easily seen that

$$w_x = u_x + v_x \quad w_y = u_y + v_y \quad w_z = u_z + v_z \quad (3-2)$$



**Figure 3-1** Addition of two vectors.

In other words, the projections of the sum are the sum of the projections, as may be seen in Figure 3-1. It is often convenient to express a vector  $\mathbf{v}$  in terms of the three unit vectors  $\mathbf{i}$ ,  $\mathbf{j}$ , and  $\mathbf{k}$  along the X, Y, and Z axes, respectively:

$$\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k} \quad (3-3)$$

Finally, it should be noted that the length of the vector  $\mathbf{v}$  is a scalar denoted by the symbol  $v$ .

There are two different products of two vectors  $\mathbf{u}$  and  $\mathbf{v}$ . The first one is a scalar, and it is known as either the inner product, scalar product, or dot product. It is defined as

$$(\mathbf{u} \cdot \mathbf{v}) = u_x v_x + u_y v_y + u_z v_z = uv \cos \theta \quad (3-4)$$

Here  $\theta$  is the angle between  $\mathbf{u}$  and  $\mathbf{v}$ . The scalar product is denoted by a dot, a pair of parentheses, or both.

The vector product, outer product, or cross product of the two vectors  $\mathbf{u}$  and  $\mathbf{v}$  is a vector  $\mathbf{w}$  defined as

$$\mathbf{w} = (u_y v_z - u_z v_y)\mathbf{i} + (u_z v_x - u_x v_z)\mathbf{j} + (u_x v_y - u_y v_x)\mathbf{k} \quad (3-5)$$

It is denoted by a cross, a pair of square brackets, or both:

$$\mathbf{w} = [\mathbf{u} \times \mathbf{v}] \quad (3-6)$$

Its length is

$$w = uv \sin \theta \quad (3-7)$$

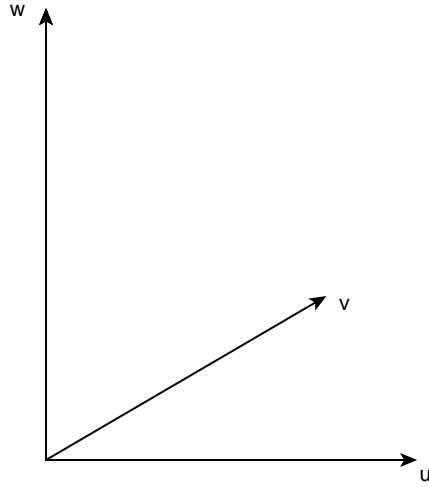
and its direction is perpendicular to the plane of the vectors  $\mathbf{u}$  and  $\mathbf{v}$  pointing in the direction of a corkscrew turning from  $\mathbf{u}$  to  $\mathbf{v}$  (see Figure 3-2).

Scalar and vector fields have played an important role in the theoretical descriptions of fluid motion and electromagnetism. They have also proved helpful in the mathematical formulation of classical and quantum mechanics, and we therefore present a brief outline of their properties.

A scalar field  $\phi(x, y, z; t)$  is simply a function of the coordinates  $x$ ,  $y$ , and  $z$  and, in addition, of the time  $t$ . By analogy, a vector field  $\mathbf{v}(x, y, z; t)$  consists of three functions  $v_x(x, y, z; t)$ ,  $v_y(x, y, z; t)$ , and  $v_z(x, y, z; t)$  of place and time. It will be assumed that the above are all continuous functions.

An important category of vector fields may be obtained as partial derivatives of scalar fields. If we define the vector  $\mathbf{v}$  as

$$\mathbf{v}(x, y, z; t) = \frac{\partial \phi}{\partial x} \mathbf{i} + \frac{\partial \phi}{\partial y} \mathbf{j} + \frac{\partial \phi}{\partial z} \mathbf{k} \quad (3-8)$$



**Figure 3-2** The vector product of two vectors.

we call  $\mathbf{v}$  the gradient of  $\phi$  and we write it as

$$\mathbf{v} = \text{grad } \phi = \nabla \phi \quad (3-9)$$

The symbol  $\nabla$  is a shorthand notation for

$$\nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \quad (3-10)$$

We define two additional quantities. The first one is a scalar,

$$\text{div } \mathbf{v} = \nabla \cdot \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \quad (3-11)$$

The second is a vector  $\mathbf{w}$ ,

$$\begin{aligned} \mathbf{w} &= \text{curl } \mathbf{v} = [\nabla \times \mathbf{v}] \\ &= \left( \frac{\partial v_y}{\partial z} - \frac{\partial v_z}{\partial y} \right) \mathbf{i} + \left( \frac{\partial v_z}{\partial x} - \frac{\partial v_x}{\partial z} \right) \mathbf{j} + \left( \frac{\partial v_x}{\partial y} - \frac{\partial v_y}{\partial x} \right) \mathbf{k} \end{aligned} \quad (3-12)$$

It may be derived from hydrodynamics or from the theory of fluid motion that

$$\frac{\partial \phi}{\partial t} + \text{div } \mathbf{v} = 0 \quad (3-13)$$

if  $\mathbf{v}$  is the gradient of  $\phi$ . It may also be shown that

$$\text{curl } \mathbf{v} = \mathbf{0} \quad (3-14)$$

If  $\mathbf{v}$  is the gradient of a scalar field, the opposite is also true.

We hope that the above brief survey is helpful to those readers who are not already familiar with the subject of vector analysis.

### III. HAMILTONIAN MECHANICS

The classical laws of motion described in Section 3.I are quite straightforward, but their application to complex systems may become rather complicated. Some of the greatest mathematicians of the eighteenth century developed new mathematical techniques to calculate the motion of the planets in our solar system. One approach was formulated by *William Rowan Hamilton* (1805–1865), the Royal Astronomer of Ireland. Eventually the Hamiltonian method played an important role in Schrödinger's wave mechanics, and we present a brief description of this technique.

For simplicity, we consider the motion of a particle with mass  $m$  in three-dimensional space that is subject at each point in space to a force  $\mathbf{F}(x, y, z)$ . It follows from the law of conservation of energy that the vector field  $\mathbf{F}(x, y, z)$  may be represented as the gradient of a scalar field  $V(x, y, z)$ :

$$\begin{aligned} \mathbf{F} &= -\text{grad } V \\ &= -\frac{\partial V}{\partial x} \mathbf{i} - \frac{\partial V}{\partial y} \mathbf{j} - \frac{\partial V}{\partial z} \mathbf{k} \end{aligned} \quad (3-15)$$

It may be shown that the energy  $E$  of the particle is given by

$$E = \frac{p^2}{2m} + V(x, y, z) \quad (3-16)$$

It has a constant value because of the conservation of energy law. In Hamiltonian mechanics it is customary to use the momentum  $p$  of the particle rather than its velocity  $v$ . The momentum is defined as

$$\mathbf{p} = m\mathbf{v} \quad (3-17)$$

Here  $m$  is the mass of the particle. We have already mentioned this definition in Eq. (1-9).

In order to formulate the equations of motion in a different way, Hamilton introduced a function  $H$  that represents the functional dependence of the energy on the coordinates and momenta of the particle:

$$H(\mathbf{r}, \mathbf{p}) = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V(x, y, z) \quad (3-18)$$

This expression is identical to Eq. (3-16), but the difference is that  $E$  represents a value, whereas the Hamiltonian function  $H$  is a function.

It may be seen that  $H$  is the sum of two terms. The first term is called the kinetic energy and is usually denoted by the symbol  $T$ ; it depends on the momentum components  $p_x$ ,  $p_y$ , and  $p_z$  but not on the coordinates. The second term is the potential energy  $V$ ; it depends on the coordinates but not on the momenta. Hamilton now proposed the following equations of motion for the particle:

$$\begin{aligned}\frac{\partial x}{\partial t} &= \frac{p_x}{m} = \frac{\partial T}{\partial p_x} = \frac{\partial H}{\partial p_x}, \quad \text{etc.} \\ \frac{\partial p_x}{\partial t} &= ma_x = F_x = -\frac{\partial V}{\partial x} = -\frac{\partial H}{\partial x}, \quad \text{etc}\end{aligned}\tag{3-19}$$

It is customary to formulate the Hamiltonian equations of motion as follows:

$$\frac{\partial r_i}{\partial t} = \frac{\partial H}{\partial p_i} \quad \frac{\partial p_i}{\partial t} = -\frac{\partial H}{\partial r_i} \quad (i = x, y, z)\tag{3-20}$$

The reader will probably notice that the Hamiltonian formalism does not seem to offer any advantages in dealing with one-particle systems. It does, however, have the advantage that it can easily be generalized. Let us, for instance, consider a system that is described by a set of generalized coordinates  $q_1, q_2, \dots, q_N$ . Each coordinate  $q_i$  is paired off with a corresponding conjugate momentum  $p_i$ . The Hamiltonian equations of motion for this system are then given by

$$\frac{\partial q_i}{\partial t} = \frac{\partial H}{\partial p_i} \quad \frac{\partial p_i}{\partial t} = -\frac{\partial H}{\partial q_i}\tag{3-21}$$

The most important reason for describing the Hamiltonian formalism here is, of course, its relation to quantum mechanics because Schrödinger used it as a basis for the mathematical formulation of his wave equation.

#### IV. THE CLASSICAL HARMONIC OSCILLATOR

We give an example of classical mechanics by describing the motion of the harmonic oscillator. It consists of a particle of mass  $m$  oscillating back and forth around the origin. The motion is one-dimensional, and the particle is subject to a force  $F$  that is proportional to the distance  $x$  between the particle and the origin.

$$F = -kx\tag{3-22}$$

At the beginning of the twentieth century, the harmonic oscillator was a favorite model for the representation of the motion of electrons in atoms and molecules.

There is no need to make use of Hamiltonian mechanics in order to describe the harmonic oscillator since its motion may simply be derived from Eq. (3-1). We may write this as

$$m \frac{d^2x}{dt^2} = -kx \quad (3-23)$$

where  $m$  is the mass of the particle. It is convenient to rewrite this equation as

$$\frac{d^2x}{dt^2} + \omega^2 x = 0 \quad (3-24)$$

where we have introduced the angular frequency  $\omega$

$$\omega^2 = \frac{k}{m} \quad (3-25)$$

The above differential equation was discussed in Eq. (2-6), and its general solution is given by

$$x(t) = B \exp(i\omega t) + C \exp(-i\omega t) \quad (3-26)$$

where  $B$  and  $C$  are two arbitrary parameters. We may simplify the solution to

$$x(t) = A \sin \omega t \quad (3-27)$$

by imposing the condition that

$$x(0) = 0 \quad (3-28)$$

The particle oscillates between the two points  $x = A$  and  $x = -A$ , and  $A$  is therefore called the amplitude of the harmonic oscillator.

The energy of the harmonic oscillator is given by

$$E = H = T + V = \frac{1}{2}(mv^2 + kx^2) = \frac{1}{2}mA^2\omega^2 \quad (3-29)$$

For given values of mass  $m$  and amplitude  $A$  the energy of the harmonic oscillator is a continuous function of the angular frequency  $\omega$ . In classical mechanics the energy  $E$  of the harmonic oscillator can therefore assume all possible values.

## V. ANGULAR MOMENTUM

An important application of both classical and quantum mechanics is the motion of a particle in a central force field where the potential function  $V$  depends only on the

distance  $r$  between the particle and the origin of the coordinate system. In this case, there is in addition to the energy  $E$  a second physical quantity that remains constant in time, namely, the angular momentum vector, which is usually denoted by the symbol  $\mathbf{M}$ . In German and in Dutch,  $\mathbf{M}$  is called the rotational momentum, which may actually be a more suitable name.

The angular momentum vector is defined as the vector product of  $\mathbf{r}$  and  $\mathbf{p}$ ,

$$\mathbf{M} = [\mathbf{r} \times \mathbf{p}] \quad (3-30)$$

or

$$M_x = yp_z - zp_y \quad M_y = zp_x - xp_z \quad M_z = xp_y - yp_x \quad (3-31)$$

It may now be shown that the time derivative of any of the three components is zero:

$$\begin{aligned} \frac{\partial M_x}{\partial t} &= \frac{\partial y}{\partial t} p_z - \frac{\partial z}{\partial t} p_y + y \frac{\partial p_z}{\partial t} - z \frac{\partial p_y}{\partial t} \\ &= yF_z - zF_y = -y \frac{\partial V}{\partial z} - z \frac{\partial V}{\partial y} = 0 \end{aligned} \quad (3-32)$$

The motion of a particle in a central force field is usually described by introducing polar coordinates  $(r, \theta, \phi)$  instead of the Cartesian coordinates  $(x, y, z)$ . We find it convenient to separate the transformation into two steps (see Figure 3-3). We define the projection of the vector  $\mathbf{r}$  on the  $xy$  plane as a vector  $\rho$ , and we then have

$$z = r \cos \theta \quad \rho = r \sin \theta \quad (3-33)$$

where  $\theta$  is the angle between the vector  $\mathbf{r}$  and the  $Z$  axis. We also have

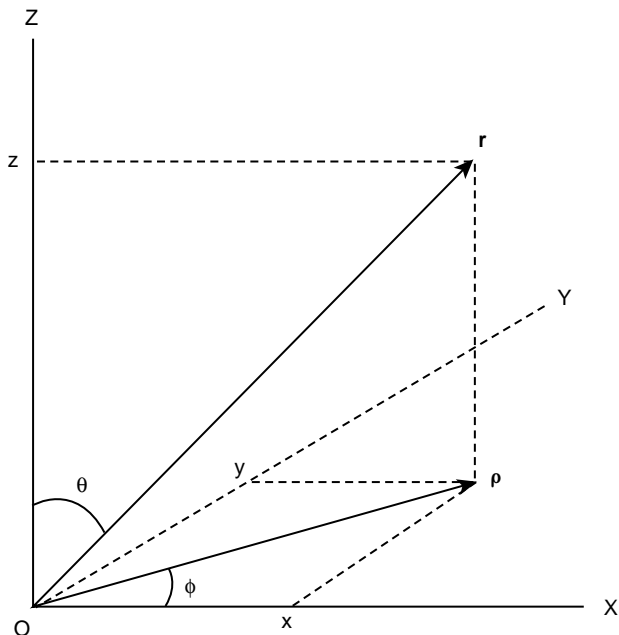
$$x = \rho \cos \phi \quad y = \rho \sin \phi \quad (3-34)$$

where  $\phi$  is the angle between  $\rho$  and the  $X$  axis. Combining the two equations (3-33) and 3-34) then gives

$$\begin{aligned} x &= r \sin \theta \cos \phi \\ y &= r \sin \theta \sin \phi \\ z &= r \cos \theta \end{aligned} \quad (3-35)$$

As an example, we will present a brief description of the Kepler problem, named after the mathematician and astronomer *Johannes Kepler* (1571–1630). The Kepler problem describes the motion of a particle in a central force field where the attractive potential  $V(r)$  is inversely proportional to the distance  $r$  between the particle





**Figure 3-3** Definition of polar coordinates.

and the origin. Examples are the motion of the moon around the Earth or the motion of the Earth around the sun. The classical description of the hydrogen atom corresponding to the motion of an electron around the hydrogen nucleus is also an example of the Kepler problem. The potential function  $V(r)$  may be represented in all of these cases as

$$V(r) = -\frac{A}{r} \quad (3-36)$$

The mathematical description of the particle motion in the Kepler problem may be derived from its two constants of the motion, namely, the energy  $E$  and the angular momentum vector  $\mathbf{M}$ . Since the vector  $\mathbf{M}$  is constant, we may choose the  $Z$  axis along its direction. The motion is then confined to the  $XY$  plane and the momentum is given by

$$\begin{aligned} M &= M_z = m \left[ x \frac{\partial y}{\partial t} - y \frac{\partial x}{\partial t} \right] \\ &= m \left[ \rho \cos \phi \left( \frac{\partial \rho}{\partial t} \sin \phi + \rho \cos \phi \frac{\partial \phi}{\partial t} \right) \right. \\ &\quad \left. - \rho \sin \phi \left( \frac{\partial \rho}{\partial t} \cos \phi - \rho \sin \phi \frac{\partial \phi}{\partial t} \right) \right] = m \rho^2 \frac{\partial \phi}{\partial t} \end{aligned} \quad (3-37)$$

where  $m$  is the mass of the particle.

The energy of the particle is given by

$$E = \frac{m}{2} \left[ \left( \frac{dx}{dt} \right)^2 + \left( \frac{dy}{dt} \right)^2 \right] - \frac{A}{r} \quad (3-38)$$

By making use of the transformation to polar coordinates (3-34) and by substituting the result (3-36) for the angular momentum, we obtain the following equations:

$$E = \frac{m}{2} \left( \frac{d\rho}{dt} \right)^2 + \frac{M^2}{2m\rho^2} - \frac{A}{r} \quad (3-39)$$

$$M = m\rho^2 \frac{d\phi}{dt}$$

It is now possible to obtain the solution of the Kepler problem by expressing the derivatives of  $\rho$  and  $\phi$  in terms of  $E$  and  $M$  and by subsequent integration. This is a rather complex procedure, and we do not describe it in detail but we will mention some of the results.

We assume that the energy  $E$  is negative, and we may then determine that in that case the particle is moving in an elliptical orbit where the relation between the radius  $r$  and the angle  $\phi$  is given by

$$\rho = \frac{M^2}{mA} \cdot \frac{1}{1 + \varepsilon \cos \phi} \quad (3-40)$$

The parameter  $\varepsilon$  is called the excentricity of the orbit, and its value depends on both  $E$  and  $M$ . An interesting feature of the motion is related to the surface area that is swept out by the radius vector while the particle moves. It may be seen from Figure 3-4 that the area  $dS$  that is swept if the polar angle  $\phi$  increases by an amount  $d\phi$  is given by

$$dS = \frac{1}{2} \rho^2 d\phi \quad (3-41)$$

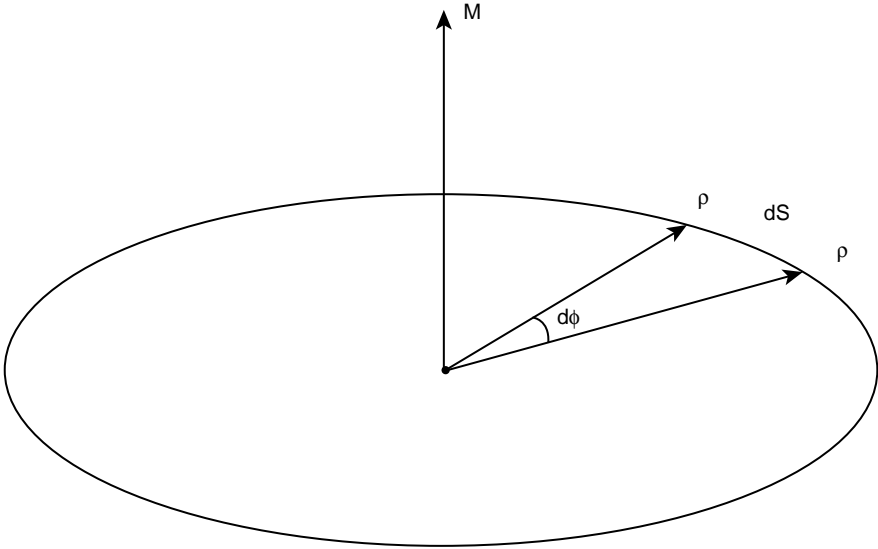
A comparison with Eq. (3-39) for the angular momentum shows that

$$\frac{dS}{dt} = \frac{M}{2m} \quad (3-42)$$

which is a constant. This result explains Kepler's second law of planetary motion, which states that in equal times the vector representing the planetary motion sweeps out equal areas.

Equation (3-42) also provides a relation between the surface area  $S$  of the ellipse and the time of revolution  $T$ , which is often called the period:

$$T = \frac{2mS}{M} \quad (3-43)$$



**Figure 3-4** Motion of a particle in Kepler's systems.

The period  $T$  may therefore be derived from  $S$ , which in turn may be derived from Eq. (3-39). This result has some historical significance because it was used by Bohr as a basis for his quantization rules for the hydrogen atom.

Since we introduced polar coordinates in this chapter, we felt that it might be appropriate to present some additional features of these coordinates in the last section.

## VI. POLAR COORDINATES

It is necessary to transform the Laplace operator of Eq. (1.16) from Cartesian to polar coordinates in order to derive the solutions of the Schrödinger equation for the hydrogen atom. We feel that this may be an appropriate occasion to present this derivation since we have already discussed some aspects of this transformation.

We again proceed in two steps. The first step is

$$\begin{aligned} x &= \rho \cos \phi \\ y &= \rho \sin \phi \end{aligned} \quad (3-44)$$

It follows that

$$\begin{aligned} \frac{\partial f}{\partial \rho} &= \frac{\partial x}{\partial \rho} \frac{\partial f}{\partial x} + \frac{\partial y}{\partial \rho} \frac{\partial f}{\partial y} = \cos \phi \frac{\partial f}{\partial x} + \sin \phi \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial \phi} &= \frac{\partial x}{\partial \phi} \frac{\partial f}{\partial x} + \frac{\partial y}{\partial \phi} \frac{\partial f}{\partial y} = -\rho \sin \phi \frac{\partial f}{\partial x} + \rho \cos \phi \frac{\partial f}{\partial y} \end{aligned} \quad (3-45)$$

Adding and subtracting these equations leads to

$$\begin{aligned}\frac{\partial f}{\partial x} &= \cos \phi \frac{\partial f}{\partial \rho} - \frac{\sin \phi}{\rho} \frac{\partial f}{\partial \phi} \\ \frac{\partial f}{\partial y} &= \sin \phi \frac{\partial f}{\partial \rho} + \frac{\cos \phi}{\rho} \frac{\partial f}{\partial \phi}\end{aligned}\tag{3-46}$$

By repeating these two differentiations and subsequently adding the results, we obtain

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial f}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \phi^2}\tag{3-47}$$

If we substitute this result into the expression for the Laplace operator, we find that

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{\partial^2 f}{\partial \rho^2} + \frac{\partial^2 f}{\partial z^2} + \frac{1}{\rho} \frac{\partial f}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \phi^2}\tag{3-48}$$

The second step of the transformation is

$$\begin{aligned}z &= r \cos \theta \\ \rho &= r \sin \theta\end{aligned}\tag{3-49}$$

analogous to Eq. (3-44). It follows immediately, by analogy with Eq. (3-47), that

$$\frac{\partial^2 f}{\partial z^2} + \frac{\partial^2 f}{\partial \rho^2} = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2}\tag{3-50}$$

Also, analogous to the second Eq. (3-46)

$$\frac{\partial f}{\partial \rho} = \sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta}\tag{3-51}$$

Finally, by combining Eqs. (3-48), (3-50), and (3-51) we obtain the desired transformation

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{\cos \theta}{r^2 \sin \theta} \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}\tag{3-52}$$

It should be noted that the above derivation is a typical examination question in advanced calculus courses. We present it here because we will need the result in a subsequent chapter and because it is an interesting exercise in calculus.

**VII. PROBLEMS****3-1** Prove that

$$\text{curl grad } \phi = 0$$

**3-2** Prove that

$$\text{div curl } \mathbf{v} = 0$$

**3-3** Prove that

$$\mathbf{u} \cdot [\mathbf{v} \times \mathbf{w}] = \mathbf{v} \cdot [\mathbf{w} \times \mathbf{u}] = \mathbf{w} \cdot [\mathbf{u} \times \mathbf{v}]$$

**3-4** A particle is confined to move on the surface of a sphere of radius  $R$  but is not subject to any other forces. What is its energy if we know that the magnitude of its angular momentum is  $M$ ?

**3-5** An ellipse may be defined as the path of a point the sum of whose distances from two fixed points (the foci) is constant (Webster's New Collegiate dictionary). Take the foci along the  $X$  axis at the position  $(c, 0)$  and  $(-c, 0)$  and define the sum of the distances as  $2a$ . Show that the definition represents then an ellipse with its major axes  $a$  and  $b$  along the  $X$  and  $Y$  axis respectively and determine  $b$ .

**3-6** A satellite describes an elliptical orbit around the earth and it is known that at its lowest point it moves 240 miles above the surface of the earth. At that time it moves through an arc of  $1^\circ$  (one degree) in 60 seconds. At a later time it moves through an arc of  $1^\circ$  in 72 seconds. How far is the satellite removed from the earth's surface at that time? Assume that the earth is a sphere with a radius of 3960 miles.

**3-7** Determine the surface area  $S$  of the ellipse corresponding to the orbit of the particle in the Kepler problem described by the equation of motion

$$\frac{d\rho}{dt} = \left[ \frac{2E}{m} + \frac{2e^2}{m\rho} - \frac{M^2}{m^2\rho^2} \right]^{1/2}$$

---

# 4

---

## WAVE MECHANICS OF A FREE PARTICLE

### I. INTRODUCTION

We mentioned in Section 1.VI that Louis de Broglie's proposal of associating a wave with the motion of a particle was one of the most spectacular advances in the development of quantum mechanics. We now use de Broglie's hypothesis to derive the mathematical description of the motion of a free particle, that is, a particle not subject to any force. This description is more complex than might be expected because we must also take Heisenberg's indeterminacy relations into account. According to Heisenberg's principle, it is not advisable to describe the motion of the particle by a single plane wave. In the latter case, the momentum of the particle is defined exactly and the position of the particle is therefore completely undetermined. As a consequence, we would be unable to make any prediction about the particle's motion.

A more realistic mathematical model is the association of particle motion with the superposition of an infinite number of plane waves. In this approach the motion of a free particle may be represented by a so-called wave packet. This makes it possible to explain the relation between particle and wave motion in a logical and consistent manner.

We first present a brief mathematical description of plane waves. Next, we give a survey of Fourier analysis that describes the superposition of plane waves. We then describe the properties of wave packets, which finally enables us to derive the wave mechanical representation of free particle motion.

## II. THE MATHEMATICS OF PLANE WAVES

A wave may be described as a vibrational motion that is propagated in a given direction. In Figure 1-1 we sketched a wave with its direction of propagation along the  $Z$  axis. If we denote the magnitude of the vibrational motion by the letter  $D$ , then it is clear that  $D(z, t)$  is a function of both the position  $z$  and the time  $t$ . Figure 1-1 shows the wave pattern at a given time  $t$  as a function of the coordinate  $z$ . It is a periodic pattern, and if we define the distance between two wave peaks as the wavelength  $\lambda$ , it is easily seen that

$$D(z, t) = D(z + \lambda, t) \quad (4-1)$$

for every value of the coordinate  $z$ . By the same token, we note that at a fixed position  $z_0$  the motion is also periodic in the time. If we define the time period by the symbol  $T$ , we have

$$D(z, t) = D(z, t + T) \quad (4-2)$$

In physics we encounter two different types of waves, transverse and longitudinal. In transverse waves the vibrational motion is perpendicular to the direction of propagation, and in longitudinal waves the vibrational motion has the same direction as the direction of propagation. The various waves that were mentioned before, such as light waves, X rays, and radio waves, are all transverse waves. Sound waves are longitudinal waves; in this case, local density variations correspond to the vibrations. Figure 1-1 represents a transverse wave; it is not easy to sketch a longitudinal wave.

A wave is usually described by either a sine or cosine function of the form

$$D(z, t) = A \sin \left[ 2\pi \left( \frac{z}{\lambda} - \frac{t}{T} \right) + \alpha \right] \quad (4-3)$$

Here  $A$  is called the amplitude of the wave and  $\alpha$  is called a phase factor; the latter is an arbitrary constant related to the choice of origin.

It is customary to replace the wavelength  $\lambda$  by its inverse, the wave number  $\sigma$ , and the period  $T$  by its inverse, the frequency  $\nu$ :

$$\sigma = \frac{1}{\lambda} \quad \nu = \frac{1}{T} \quad (4-4)$$

Equation (4-3) then becomes

$$D(z, t) = A \sin[2\pi(\sigma z - \nu t) + \alpha] \quad (4-5)$$

It is also convenient to represent the wave as a complex quantity

$$\phi(z, t) = A \exp[2\pi i(\sigma z - \nu t)] \quad (4-6)$$

The phase factor  $\alpha$  has now been incorporated into the coefficient  $A$ .

We shall see that from a mathematical point of view, Eq. (4-6) is the most convenient to use. It should be realized, though, that physical observables are always real quantities; they may be obtained as either the real or the imaginary parts of Eq. (4-6).

We should realize that Eq. (4-6) is not limited to one dimension. We may assume that it represents a plane wave in three dimensions, with its direction of propagation coinciding with the Z axis and  $\phi(z, t)$  constant in any XY plane. We may generalize the description by introducing the wave vector  $\sigma$ , which points in the direction of propagation of the wave and has a magnitude that is equal to the inverse of the wavelength  $\lambda$ . The wave is then represented by

$$\begin{aligned}\phi(\mathbf{r}, t) &= A \exp[2\pi i(\sigma \cdot \mathbf{r} - vt)] \\ &= A \exp[2\pi i(\sigma_x x + \sigma_y y + \sigma_z z - vt)]\end{aligned}\quad (4-7)$$

The definition of  $\phi$  depends on its application. In the case of light waves it is related to electric and magnetic field strengths. We are, of course, interested in de Broglie waves, and in our case  $\phi$  is related to the motion of a free particle.

### III. THE SCHRÖDINGER EQUATION OF A FREE PARTICLE

In Chapter 1, we described how Schrödinger made a critical contribution to the development of quantum mechanics by showing that de Broglie's waves may be represented as solutions of a differential equation. The derivation of the equation is based on Planck's expression for the energy quantization and on de Broglie's relation (1-14) between momentum and wavelength, which we may write as

$$E = h\nu \quad \mathbf{p} = h\sigma \quad (4-8)$$

The potential energy of a free particle is equal to zero, and we therefore have

$$E = H(\mathbf{r}, \mathbf{p}) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) \quad (4-9)$$

according to Eq. (3-18). Substitution of Eq. (4-8) gives

$$h\nu = \frac{h^2}{2m}(\sigma_x^2 + \sigma_y^2 + \sigma_z^2) \quad (4-10)$$

This equation is known as the dispersion relation of the de Broglie waves.

It is easily seen that the plane waves of Eq. (4-7) satisfy the following relations:

$$\frac{\partial \phi}{\partial t} = -2\pi i\nu\phi \quad \frac{\partial \phi}{\partial x} = 2\pi i\sigma_x\phi \quad \text{etc.} \quad (4-11)$$



Substitution into Eq. (4-10) leads to the following differential equation:

$$-\frac{\hbar^2}{2m}\Delta\phi = -\frac{\hbar}{i}\frac{\partial\phi}{\partial t} \quad (4-12)$$

This is known as the free particle Schrödinger equation.

Any plane wave defined by Eq. (4-7) is a solution of the Schrödinger equation (4-12), and the general solution of the latter equation is a generalized linear combination of all its solutions, which we may write as

$$\Psi(\mathbf{r}, t) = \int A(\boldsymbol{\sigma}) \exp[2\pi i(\boldsymbol{\sigma} \cdot \mathbf{r} - vt)] d\boldsymbol{\sigma} \quad (4-13)$$

This is the general expression for a de Broglie wave associated with the motion of a free particle if we substitute the dispersion relation (4-10) for the frequency.

Let us now proceed to a discussion of the time evolution of the generalized de Broglie wave of Eq. (4-13). For simplicity, we consider the one-dimensional case since its generalization to three dimensions is trivial. Our discussion is based on the Fourier integral theorem, which is given by

$$f(x) = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} f(t) \exp[-2\pi i u(t-x)] dt \quad (4-14)$$

and which constitutes the basis of Fourier analysis.

Equation (4-14) may be separated into two equations:

$$F(u) = \int_{-\infty}^{\infty} f(t) \exp(-2\pi i ut) dt \quad (4-15a)$$

$$f(x) = \int_{-\infty}^{\infty} F(u) \exp(2\pi i ux) du \quad (4-15b)$$

Each of these equations represents a Fourier integral transform, a specific type of integral transform.

An integral transform describes the transformation of a function  $f$  into a different function  $g$  by means of an integration:

$$g(y) = \int f(x) K(x, y) dx \quad (4-16)$$

The specific type of integral transform depends, of course, on the specific form of the function  $K(x, y)$ . Integral transforms are useful tools in mathematical analysis. They have been extensively tabulated, so that not only the transforms but also the so-called inverse transforms (where the function  $f$  is derived from the function  $g$ ) are widely known. The Fourier transform of Eq. (4-15) is characterized by a unique

feature; the transform and its inverse are represented by the same integration, apart from a different sign in the exponential.

The mathematical procedure for deriving the time evolution of the de Broglie wave is now easily derived by making use of the Fourier integral theorem. At time  $t = 0$  we have, according to Eq. (4-13),

$$\Psi(x, 0) = \int A(\sigma) \exp[2\pi i \sigma x] d\sigma \quad (4-17)$$

The inverse transform, as described by Eq. (4-15a), is

$$A(\sigma) = \int \Psi(x, 0) \exp[-2\pi i \sigma x] dx \quad (4-18)$$

Substitution of this result together with the dispersion relation (4-10) into Eq. (4-13) gives

$$\Psi(x, t) = \int_{-\infty}^{\infty} A(\sigma) \exp[2\pi i(\sigma x - \sigma^2 ht/2m)] d\sigma \quad (4-19)$$

It follows that any initial function  $\Psi(x, 0)$  may be represented as a superposition of plane waves and that the time dependence of any such function may be predicted by straightforward integration.

We will illustrate the above mathematical technique later by describing an example that was first presented by Heisenberg in order to support the indeterminacy relations.

#### IV. THE INTERPRETATION OF THE WAVE FUNCTION

In 1926 Max Born published an important paper where he used the newly developed quantum mechanics to describe the scattering of a beam of electrons by an atom or, to be more precise, by a central potential field that decreases faster than the inverse of the distance. Born proposed that the probability of finding a scattered electron in a given direction defined by a small element of solid angle is related to the value of the wave function at that solid angle. In a footnote added in proof, Born corrected this statement by relating the probability to the square of the wave function. He failed to recognize the possibility that the wave function could be a complex quantity. Since the probability must always be positive, the true answer is that it is determined by the product of the wave function and its complex conjugate.

It became universally accepted that the probability  $P(x, y, z; t)$  of finding a particle in a volume element  $dx dy dz$  surrounding a point  $(x, y, z)$  is given by the product of the wave function  $\Psi$  and its complex conjugate  $\Psi^*$  at that point:

$$P(x, y, z; t) = \Psi(x, y, z; t) \cdot \Psi^*(x, y, z; t) dx dy dz \quad (4-20)$$

Since the total probability of finding the particle is equal to unity, we must impose the condition

$$\int \int \int \Psi(x, y, z; t) \cdot \Psi^*(x, y, z; t) dx dy dz = 1 \quad (4-21)$$

It is interesting to note that initially Born received little credit for the statistical interpretation of the wave function. His idea was universally adopted, but most prominent scientists were not aware of the fact that Born was the first to propose it. Eventually Born received credit for his probabilistic interpretation of the wave function since it was cited when he was awarded the Nobel Prize for physics in 1954.

Even though the wave function offers only a statistical representation of the motion of a particle, it determines the so-called expectation value of a physical observable that constitutes a well-defined exact prediction of the particle's behavior. We again limit the following definitions to the one-dimensional case since generalizations to more dimensions are obvious.

The expectation value  $\langle x \rangle$  of the position of a particle is defined as the integral

$$\langle x \rangle = \int \Psi^*(x, t) x \Psi(x, t) dx \quad (4-22)$$

Here again we impose the condition

$$\int \Psi^*(x, t) \Psi(x, t) dx = 1 \quad (4-23)$$

analogous to Eq. (4-21). The expectation value of a function  $f(x)$  of the coordinate  $x$  is defined in a similar way as

$$\langle f(x) \rangle = \int \Psi^*(x, t) f(x) \Psi(x, t) dx \quad (4-24)$$

The expectation value  $\langle \sigma \rangle$  of the wave vector is related to the function  $A(\sigma)$  and is defined as

$$\langle \sigma \rangle = \int A^*(\sigma) \sigma A(\sigma) d\sigma \quad (4-25)$$

The expectation value of a function  $g(\sigma)$  of  $\sigma$  is defined in a similar way as

$$\langle g(\sigma) \rangle = \int A^*(\sigma) g(\sigma) A(\sigma) d\sigma \quad (4-26)$$

We must again impose the condition

$$\int A^*(\sigma)A(\sigma)d\sigma = 1 \quad (4-27)$$

It should be noted that the expectation value  $\langle \sigma \rangle$  is time independent. This is consistent with classical physics since the momentum and consequently the wave vector of a particle that is not subject to outside forces should not be dependent on the time.

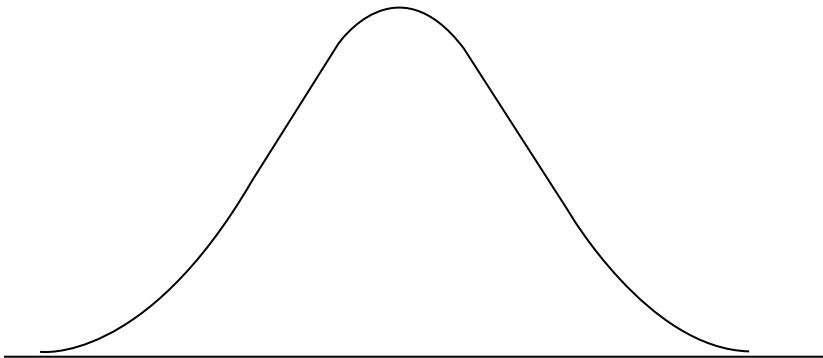
## V. WAVE PACKETS

We mentioned in Section 4.I that the preferred mathematical model for representing the motion of a free particle is by means of a wave packet. This may be defined as a superposition of waves having the same or almost the same phase in the vicinity of a given point  $\mathbf{r}_0$  and different random phases everywhere else. The wave function of a wave packet is therefore different from zero in the vicinity of  $\mathbf{r}_0$  and very small or zero everywhere else (see Figure 4-1). We will first present some general features of wave packets, and then we will discuss a specific example.

The velocity of propagation of a single wave is called its phase velocity. Figure 1-1 shows that such a wave travels distance  $\lambda$  during time  $T$ , and its phase velocity  $v$  is therefore given by

$$v = \frac{\lambda}{T} = \lambda\nu \quad (4-28)$$

A wave packet is also characterized by its group velocity, that is, the velocity  $u$  at which the maximum of the wave packet moves along. The group velocity may also



**Figure 4-1** Sketch of a typical wave packet.

be defined as the velocity of the expectation value of the position  $\langle \mathbf{r} \rangle$ ,

$$\mathbf{u} = \frac{d\langle \mathbf{r} \rangle}{dt} \quad (4-29)$$

In order to determine the group velocity, we note that at given time  $t_0$  the various waves of the wave packet all have the same phase at the point  $\mathbf{r}_0$ , and the latter point is therefore characterized by the condition

$$x_0 d\sigma_x + y_0 d\sigma_y + z_0 d\sigma_z - t_0 dv = 0 \quad (4-30)$$

A small time interval  $dt$  later the maximum has moved to  $\mathbf{r}_0 + d\mathbf{r}$ , and at this new maximum position we have

$$(x_0 + dx)d\sigma_x + (y_0 + dy)d\sigma_y + (z_0 + dz)d\sigma_z - (t_0 + dt)dv = 0 \quad (4-31)$$

since

$$u_x = \frac{dx}{dt} \quad u_y = \frac{dy}{dt} \quad u_z = \frac{dz}{dt} \quad (4-32)$$

We find by subtracting Eq. (4-30) from Eq. (4-31) that

$$u_x d\sigma_x + u_y d\sigma_y + u_z d\sigma_z - dv = 0 \quad (4-33)$$

We should realize that the frequency is related to the wave vector  $\sigma$  according to the dispersion relation; we therefore have

$$dv = \frac{\partial v}{\partial \sigma_x} d\sigma_x + \frac{\partial v}{\partial \sigma_y} d\sigma_y + \frac{\partial v}{\partial \sigma_z} d\sigma_z \quad (4-34)$$

The group velocity of the wave packet is therefore given by

$$u_x = \frac{\partial v}{\partial \sigma_x} \quad u_y = \frac{\partial v}{\partial \sigma_y} \quad u_z = \frac{\partial v}{\partial \sigma_z} \quad (4-35)$$

It may be recalled that the dispersion relation was given by Eq. (4-10)

$$v = \frac{h}{2m} (\sigma_x^2 + \sigma_y^2 + \sigma_z^2) \quad (4-36)$$

so that

$$u_x = \frac{h\sigma_x}{m} = \frac{p_x}{m} \quad \text{etc.} \quad (4-37)$$

In other words, the group velocity of the wave packet is equal to the classical velocity of the free particle.

In order to illustrate the properties of wave packets, we discuss a specific one-dimensional case that was analyzed by Heisenberg to show the connection with the indeterminacy relations. In the mathematical analysis, Heisenberg made use of the following two definite integral results:

$$I = \int_{-\infty}^{\infty} \exp(-ax^2) dx = (\pi/a)^{1/2} \quad (4-38)$$

$$J = \int_{-\infty}^{\infty} x^2 \exp(-ax^2) dx = (I/2a)$$

Heisenberg calculated the behavior of a one-dimensional wave packet whose wave function  $\Psi(x; 0)$  at the initial time  $t = 0$  is given by

$$\Psi(x; 0) = (2\tau)^{1/4} \exp(-\pi\tau x^2 + 2\pi i\sigma_0 x) \quad (4-39)$$

The probability density corresponding to this wave packet is a Gaussian function given by

$$\Psi^*(x; 0) \cdot \Psi(x; 0) = (2\tau)^{1/2} \exp(-2\pi\tau x^2) \quad (4-40)$$

We have sketched this Gaussian function in Figure 4-1.

Using Eq. (4-38), it is easily verified that the expectation values of  $x$  and of  $x^2$  at the initial time  $t = 0$  are given by

$$\begin{aligned} \langle x \rangle_0 &= 0 \\ \langle x^2 \rangle_0 &= 1/4\pi\tau \end{aligned} \quad (4-41)$$

The amplitude function  $A(\sigma)$  may now be derived by substituting Eq. (4-39) into Eq. (4-18):

$$\begin{aligned} A(\sigma) &= \int_{-\infty}^{\infty} \Psi(x; 0) \exp(-2\pi i\sigma x) dx \\ &= (2\tau)^{1/4} \int_{-\infty}^{\infty} \exp[-\pi\tau x^2 - 2\pi i(\sigma - \sigma_0)x] dx \end{aligned} \quad (4-42)$$

After some rearrangements and by making use of Eq. (4-38), we find that

$$A(\sigma) = (2\tau)^{1/4} \exp[-\pi(\sigma - \sigma_0)^2/\tau] \quad (4-43)$$

This function is again a Gaussian with its maximum at the point  $\sigma = \sigma_0$ , and the expectation value of  $\sigma$  is obvious, as given by

$$\langle \sigma \rangle = \int_{-\infty}^{\infty} A^*(\sigma) \sigma A(\sigma) d\sigma = \sigma_0 \quad (4-44)$$

The expectation value of  $(\sigma - \sigma_0)^2$  is given by

$$\langle (\sigma - \sigma_0)^2 \rangle = \int_{-\infty}^{\infty} A^*(\sigma) (\sigma - \sigma_0)^2 A(\sigma) d\sigma = \tau/4\pi \quad (4-45)$$

In order to relate the above results to the uncertainty relations, we note that the product of the uncertainty  $\Delta x$  in the coordinate and the uncertainty  $\Delta \sigma$  in the wave vector is obtained by multiplying Eqs. (4-41) and (4-45):

$$\Delta x \cdot \Delta \sigma = (1/4\pi\tau)^{1/2} (\tau/4\pi)^{1/2} = 1/4\pi \quad (4-46)$$

It follows that the product of the uncertainties  $\Delta x$  and  $\Delta p$  of the coordinate and its momentum is given by

$$\Delta x \cdot \Delta p = h/4\pi = \hbar/2 \quad (4-47)$$

which is exactly half of the value prescribed by Heisenberg's uncertainty principle. We mentioned before that it is possible to construct an ideal situation where the product of  $\Delta x$  and  $\Delta p$  is smaller than  $\hbar$ . This does not invalidate the uncertainty principle since the latter was derived by comparing the matrix representations of the coordinates and their conjugate momenta.

The amplitude function  $A(\sigma)$  is time independent, but the time-dependent wave function  $\Psi(x, t)$  has a rather complicated form. It is obtained by substituting Eq. (4-43) into Eq. (4-19):

$$\Psi(x, t) = \left(\frac{2}{\tau}\right)^{1/4} \int_{-\infty}^{\infty} \exp\left[-\frac{\pi(\sigma - \sigma_0)^2}{\tau} + 2\pi i \sigma x - \frac{\pi i h \sigma^2 t}{m}\right] d\sigma \quad (4-48)$$

The integral may be evaluated by making a few substitutions and making use of Eq. (4-38). The result of the integration is rather complex, and we prefer to list instead the probability density since it may be presented in a much more simple form

$$P(x; t) = \Psi^*(x; t) \cdot \Psi(x; t) = (2\tau')^{1/2} \exp[-2\pi\tau'[x - v_0 t]^2] \quad (4-49)$$

where

$$\tau' = \tau \left(1 + \frac{h^2 \tau^2 t^2}{m^2}\right)^{-1} \quad v_0 = \frac{h\sigma_0}{m} \quad (4-50)$$

It follows that the wave packet moves with the classical velocity  $v_0$ , which is equal to the group velocity of the wave packet. Another interesting result is the increase in the width of the wave packet since

$$\begin{aligned} \langle x \rangle_t &= (v_0 t) \\ (\Delta x)_t^2 &= \langle (x - v_0 t)^2 \rangle = \frac{1}{4\pi\tau'} = \frac{1}{4\pi\tau} \left( 1 + \frac{h^2\tau^2 t^2}{m^2} \right) \end{aligned} \quad (4-51)$$

It follows that both  $\Delta x$  and the product  $\Delta x \cdot \Delta p$  increase as a function of time, so that the predictions of the future positions of the particle become less and less precise as time progresses. In this respect, wave mechanics bears some resemblance to meteorology, where short-term predictions may be fairly reliable, whereas long-term predictions become less and less reliable with increasing time.

## VI. CONCLUDING REMARKS

The wave mechanical description of free particle motion was based on the general mathematical expression (4-8) for de Broglie waves and on the concept of a wave packet. The mathematical analysis of the time evolution of a wave packet made use of Fourier analysis, in particular the Fourier integral theorem (4-14), and the physical interpretation of the wave function was based on ideas that were first proposed by Born and described in Section 4.IV. We illustrated in Section 4.V that all results are consistent with classical mechanics and with the Heisenberg indeterminacy principle.

We also showed that the de Broglie wave function is the solution of a differential equation, namely, the free particle Schrödinger equation (4-12). It should, of course, be noted that our mathematical analysis was based on our knowledge of the general expression (4-8) of the Broglie waves, and there was no need to solve the Schrödinger equation. We basically derived the Schrödinger equation from its solution, not the reverse.

We shall see that the situation is very different for bound particles. Here Schrödinger used the same argument as for a free particle, but he incorporated the potential energy into the Hamiltonian in order to arrive at the Schrödinger equation for a particle moving in a potential field. We will present these ideas in the next chapter.

## VII. PROBLEMS

- 4-1 Calculate the energy and momentum of an X-ray quantum with a wave length of 1 Ångstrom unit. What is the velocity of an electron having a momentum equal to that of the above X-ray quantum?



- 4-2** Derive the results for the integrals  $I$  and  $J$  of Eq. (4-38) by using polar coordinates.
- 4-3** Derive that for any wave packet represented by a wave function  $\Psi(x; t)$  the expectation value

$$\langle x \rangle_t = \langle \Psi(x; t) | x | \Psi(x; t) \rangle$$

is given by

$$\langle x \rangle_t - \langle x \rangle_{t=0} = (ht/m) \langle A(\sigma) | \sigma | A(\sigma) \rangle$$

- 4-4** Derive an expression for the group velocity of an arbitrary wave packet. Show that for a de Broglie wave packet this result is consistent with the classical velocity of the particle.

---

# 5

---

## THE SCHRÖDINGER EQUATION

### I. INTRODUCTION

We showed in the preceding chapter that the de Broglie wave function of a free particle may be represented as the general solution of a partial differential equation known as the free-particle Schrödinger equation. A free particle is defined here as a particle that is not subject to any force. It is therefore characterized by a potential energy that is equal to zero and by a positive total energy  $E$ .

The free-particle Schrödinger equation was derived by making use of the properties of de Broglie waves that were given by Eq. (4-11) and that may also be written as

$$E\phi = h\nu\phi = -\frac{h}{2\pi i} \frac{\partial\phi}{\partial t} \quad (5-1)$$

and as

$$p_x\phi = h\sigma_x\phi = \frac{h}{2\pi i} \frac{\partial\phi}{\partial x}, \quad \text{etc.} \quad (5-2)$$

The Schrödinger equation is then easily derived by substituting the above equations into the expression for energy:

$$E = H(\mathbf{r}, \mathbf{p}) = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) \quad (5-3)$$

The Schrödinger equation for a nonfree or bound particle may now be obtained by including a nonzero potential function  $V(x, y, z)$  in an analogous mathematical derivation. We now have

$$E = H(\mathbf{r}, \mathbf{p}) = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V(x, y, z) \quad (5-4)$$

By again substituting Eqs. (5-1) and (5-2), we obtain the following differential equation:

$$-\frac{\hbar^2}{4\pi^2 m} \left( \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} \right) + V(x, y, z)\phi = -\frac{\hbar}{2\pi i} \frac{\partial \phi}{\partial t} \quad (5-5)$$

This is known as the time-dependent Schrödinger equation. It is usually simplified by replacing Planck's constant by the symbol  $\hbar = h/2\pi$  and by introducing the Laplace operator  $\Delta$  of Eq. (1-17). The time-dependent Schrödinger equation then becomes identical to the equation

$$-\frac{\hbar^2}{2m} \Delta \phi + V\phi = -\frac{\hbar}{i} \frac{\partial \phi}{\partial t} \quad (1-16b)$$

which was mentioned in Chapter 1.

The time-dependent Schrödinger equation may be transformed into a corresponding time-independent equation by making use of Bohr's concept of stationary states with well-defined energies. A stationary state has an energy  $E$  and consequently a well-defined frequency  $\nu = E/h$ . Its wave function  $\phi(x, y, z; t)$  may therefore be represented as

$$\phi(x, y, z; t) = \psi(x, y, z) \exp(-iEt/\hbar) \quad (5-6)$$

Substitution into the time-dependent Eq. (1-16b) gives the time-independent Schrödinger equation:

$$-\frac{\hbar^2}{2m} \Delta \psi + V\psi = E\psi \quad (5-7)$$

The solutions of the time-independent Schrödinger equation contain energy  $E$  as a parameter. It now appears that, in general, those solutions are not acceptable from a physical point of view. Acceptable solutions are obtained only for specific discrete values of the energy parameter  $E$  that are known as the eigenvalues  $E_n$  of the equation. These eigenvalues correspond to the stationary states that had previously been proposed by Bohr. The corresponding solutions of the equation are known as eigenfunctions.

The Schrödinger equation may therefore be classified as a differential equation with boundary conditions. This category of equations had been studied extensively

by mathematicians because they are encountered in heat conduction, vibrating surfaces, diffusion, and other conditions. Eigenvalue problems are also related to operators, a subject that we will now discuss.

## II. OPERATORS

An operator may be described as a procedure or a command that has the effect of transforming one function  $f$  into a different function  $g$ . There is a large variety of operators. We may, for example, define the operator  $Sq$  as a command to take the square of a function, the operator  $Iv$  as a command to take the inverse of a function, and so on. The integral transforms described in Section 4.III may also be considered a class of operators.

The operators that play a role in the mathematical formulation of quantum mechanics are much more restrictive in nature. They are either differential or multiplicative operators or a combination of these two types. They are also linear operators, which are defined by the relation

$$\Lambda(f + g) = \Lambda f + \Lambda g \quad (5-8)$$

It should be noted that operators are often denoted by capital Greek or Roman letters.

An important class of operators are the Hermitian operators. In order to define these operators, we introduce the following notation:

$$\begin{aligned} \langle f | \Lambda | g \rangle &= \int f^* \Lambda g \, dq \\ \langle f | g \rangle &= \int f^* g \, dq \end{aligned} \quad (5-9)$$

where the integration is to be performed over all relevant coordinates in their appropriate intervals. A Hermitian operator  $H$  is now defined by the following relation:

$$\langle f | H | g \rangle = \langle g | H | f \rangle^* \quad (5-10)$$

which we may also write as

$$\int f^* H g \, dq = \int g H^* f^* \, dq \quad (5-11)$$

Hermitian operators are important because one of the present fundamental rules of quantum mechanics is that every physical observable may be represented by a Hermitian operator.

It is now possible to rewrite the Schrödinger equation (5-7) in terms of an operator equation by introducing the Hamiltonian operator  $H_{\text{op}}$ :

$$H_{\text{op}} = -\frac{\hbar^2}{2m}\Delta + V \quad (5-12)$$

The Laplace operator  $\Delta$  is a differential operator and  $V$  is a multiplicative operator. Both of these operators are Hermitian, and  $H_{\text{op}}$  is therefore also Hermitian. The Schrödinger equation (5-7) may now be written in the form

$$H_{\text{op}}\Psi = E\Psi \quad (5-13)$$

The eigenvalues  $E_n$  and the corresponding eigenfunctions  $\Psi_n$  may then be defined by the equation

$$H_{\text{op}}\Psi_n = E_n\Psi_n \quad (5-14)$$

as the eigenvalues and eigenfunctions of the Hamiltonian operator  $H_{\text{op}}$ .

This equation bears some resemblance to Eq. (2-64), which describes the eigenvalues and eigenvectors of a matrix. It appears that there is in fact a relation between the eigenvalue problems of operators and those of matrices, and we will discuss how the one eigenvalue problem can be transformed into the other. This transformation was used to demonstrate the equivalence of the Schrödinger equation and matrix mechanics.

In quantum mechanics, Hermitian operators play an important role because in the axiomatic presentation of the subject, one of the basis assumptions is that every physical observable can be represented by a Hermitian operator. It is therefore important to present some properties of the eigenvalues and eigenfunctions of Hermitian operators.

It is easily shown that the eigenvalues  $h_n$  of a Hermitian operator  $H$  are real. We write

$$H\Psi_n = h_n\Psi_n \quad (5-15)$$

Multiplication on the left by  $\Psi_n^*$  and subsequent integration gives

$$\langle\Psi_n|H|\Psi_n\rangle = h_n\langle\Psi_n|\Psi_n\rangle \quad (5-16)$$

The complex conjugate of this equation is

$$\langle\Psi_n|H|\Psi_n\rangle^* = h_n^*\langle\Psi_n|\Psi_n\rangle^* \quad (5-17)$$

Subtracting Eq. (5-17) from Eq. (5-16) gives

$$(h_n - h_n^*)\langle\Psi_n|\Psi_n\rangle = 0 \quad (5-18)$$

since the left-hand sides of the two equations are equal due to the fact that the operator  $H$  is Hermitian. It follows that  $h_n$  must be real since

$$h_n = h_n^* \quad (5-19)$$

We now consider two eigenfunctions,  $\Psi_n$  and  $\Psi_m$ , belonging to two different eigenvalues,  $h_n$  and  $h_m$ . We have

$$H\Psi_n = h_n\Psi_n \quad (H\Psi_m)^* = h_m\Psi_m^* \quad (5-20)$$

We also have

$$\begin{aligned} \langle \Psi_m | H | \Psi_n \rangle &= h_n \langle \Psi_m | \Psi_n \rangle \\ \langle \Psi_n | H | \Psi_m \rangle^* &= h_m \langle \Psi_n | \Psi_m \rangle^* \end{aligned} \quad (5-21)$$

Again, since  $H$  is a Hermitian operator, the left-hand sides of the two equations are equal and subtraction gives

$$(h_n - h_m) \langle \Psi_m | \Psi_n \rangle = 0 \quad (5-22)$$

or

$$\langle \Psi_m | \Psi_n \rangle = 0 \quad (5-23)$$

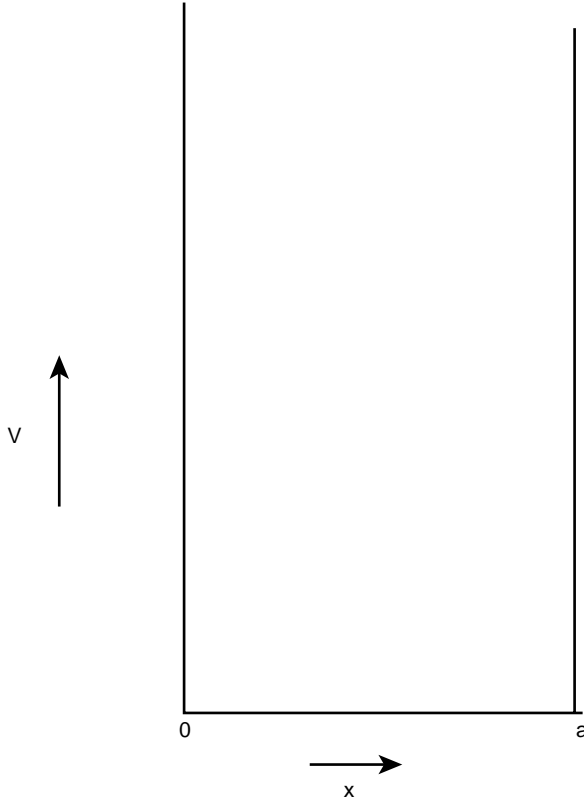
since  $h_n$  and  $h_m$  are different. We say that the two functions  $\Psi_n$  and  $\Psi_m$  are orthogonal when the integral (5-23) is zero.

### III. THE PARTICLE IN A BOX

Now that we have discussed some general features of operators and their eigenvalues and eigenfunctions, we can proceed to a practical application of the Schrödinger equation, namely, the particle in a box. This is considered the simplest problem in quantum mechanics, especially if we consider the one-dimensional case. Here the coordinate  $x$  of the particle is confined to a finite area defined by  $0 \leq x \leq a$  (see Figure 5-1).

We may represent the motion of this particle by a potential function that is zero in the region I where the particle is allowed to move and infinite outside this region; in other words

$$\begin{aligned} V(x) &= 0 & 0 \leq x \leq a & \quad \text{(I)} \\ V(x) &= \infty & x < 0 & \quad \text{(II)} \\ V(x) &= \infty & x > a & \quad \text{(III)} \end{aligned} \quad (5-24)$$



**Figure 5-1** Potential function of a particle in a one-dimensional box.

The one-dimensional Schrödinger equation is

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \phi}{\partial x^2} + V(x)\phi = E\phi \tag{5-25}$$

In region I, the potential function  $V(x)$  is zero and the Schrödinger equation reduces to

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \phi}{\partial x^2} = E\phi \tag{5-26}$$

We may assume that the energy  $E$  of the particle is positive in region I, and we may therefore write Eq. (5-26) as

$$\frac{\partial^2 \phi}{\partial x^2} = -k^2 \phi \quad k^2 = \frac{2mE}{\hbar^2} \tag{5-27}$$

We discussed this differential equation in Chapter 2. According to Eq. (2-6), its general solution is given by

$$\phi(x) = A \exp(ikx) + B \exp(-ikx) \quad (5-28)$$

Since the particle is confined to region I ( $0 \leq x \leq a$ ), the wave function  $\phi(x)$  must be zero outside this region:

$$\begin{aligned} \phi(x) &= 0 & x < 0 \\ \phi(x) &= 0 & x > a \end{aligned} \quad (5-29)$$

At this stage, we must impose suitable boundary conditions on the solution of the Schrödinger equation. In the present case, we require the wave function to be continuous everywhere, in particular at the points  $x = 0$  and  $x = a$ . This condition reduces to

$$\phi(0) = \phi(a) = 0 \quad (5-30)$$

The first boundary condition at  $x = 0$  gives

$$\phi(0) = A + B = 0 \quad (5-31)$$

The eigenfunction may now be written as

$$\begin{aligned} \phi(x) &= A[\exp(ikx) - \exp(-ikx)] \\ &= C \sin kx \end{aligned} \quad (5-32)$$

since  $\phi(x)$  may be multiplied by an arbitrary constant.

The second boundary condition is

$$\phi(a) = C \sin ka = 0 \quad (5-33)$$

which is satisfied if

$$ka = n \cdot \pi \quad n = 1, 2, 3, 4, \dots \quad (5-34)$$

The energy eigenvalues of the particle in a box are obtained by combining Eqs. (5-27) and (5-34):

$$E_n = \frac{\hbar^2 n^2}{8ma^2} \quad n = 1, 2, 3, \dots \quad (5-35)$$

The corresponding eigenfunctions  $\phi_n$  are derived by substituting Eq. (5-34) into Eq. (5-32):

$$\phi = C \sin(n\pi x/a) \quad (5-36)$$



The value of the constant  $C$  is determined by the normalization condition

$$\begin{aligned}\int_0^a f_n(x)f_n(x)dx &= C C^* \int_0^a \sin^2(\pi nx/a)dx \\ &= (a/2)C C^* = 1\end{aligned}\tag{5-37}$$

It is worth noting that this condition determines the absolute value of the constant  $C$  but not its argument, so that  $C$  is given by

$$C = \sqrt{2/a} \exp(i\gamma)\tag{5-38}$$

where  $\gamma$  is an arbitrary phase factor.

We shall see that subsequent applications require that not only the wave function but also its derivative is continuous everywhere. In the present case of a particle in a box, we imposed the continuity condition only on the wave function itself and not on its derivative. In fact, it is easily verified that the derivative of the eigenfunctions  $f_n$  is not continuous at the points  $x = 0$  and  $x = a$ .

The above example has an unusual feature compared to all other cases: the potential function  $V(x)$  is not only discontinuous at the points  $x = 0$  and  $x = a$  but also has an infinite discontinuity at these points. It can be shown that in such a case the derivative of the wave function is not required to be continuous at the points where the potential has a discontinuity of infinite magnitude. It is remarkable that the first example that we discuss and that we selected because of its mathematical simplicity exhibits an unusual feature that distinguishes it from the other cases that we will present.

#### IV. CONCLUDING REMARKS

The quantum mechanical description of a bound particle is obtained by first solving the time-independent Schrödinger equation (5-7). The general solution  $\Psi(x; E)$  of this equation contains the energy  $E$  as an adjustable parameter. It is then necessary to impose the condition that the function  $\Psi(x; E)$  corresponds to descriptions of the system that are physically realistic. In many situations, this condition would require the wave function to be normalizable; that is, the integral of the product of the wave function and its complex conjugate is finite. In other cases, the function and its derivative are required to be continuous at every point in space or to be single-valued.

It has been found that in general, only specific solutions of the Schrödinger equation are suitable for a realistic physical description of the system. These solutions are characterized by specific discrete values of the energy parameter  $E$ . These values determine the stationary states of the system. In mathematics such discrete values of the parameter are called the eigenvalues of the differential equation, but in physics they are known as the energy eigenvalues of the system. The wave functions that correspond to the eigenvalues are called the eigenfunctions of the system, and they represent all physical properties of the system.

**V. PROBLEMS**

- 5-1** Prove that the Laplace operator  $\Delta$  is Hermitian.
- 5-2** Prove that the angular momentum operator  $M^2$  is Hermitian.
- 5-3** Derive the eigenvalues and eigenfunctions of a particle in a box when its potential is given by

$$\begin{aligned} V(x) &= 0 & -a/2 \leq x \leq a/2 \\ V(x) &= \infty & x < -a/2 \text{ and } x > a/2 \end{aligned}$$

- 5-4** Show explicitly that the eigenfunctions of a particle in a box are all orthogonal.

---

# 6

---

## APPLICATIONS

### I. INTRODUCTION

Much of the study of quantum mechanics involves solving the Schrödinger equation for various systems. The standard cases that are presented in most textbooks are (1) the particle in a box, (2) the particle in a finite box, (3) a particle moving through a potential barrier (tunneling), (4) the harmonic oscillator, (5) the rigid rotor, and (6) the hydrogen atom. It should be noted that the majority of these systems were described by Schrödinger in his early papers.

The hydrogen atom is a special case of a particle in a three-dimensional central force field where the potential function  $V(r)$  depends only on the distance  $r$  between the particle and the origin. In order to study such systems, it is advantageous to introduce the angular momentum and to derive its eigenvalues and eigenfunctions first. The latter are closely related to the solutions of the rigid rotor. For these reasons, we decided to discuss the hydrogen atom and the rigid rotor in subsequent chapters, while the other applications are presented in this chapter.

### II. A PARTICLE IN A FINITE BOX

We define the potential function of a one-dimensional finite box as sketched in Figure 6-1:

$$V(x) = U \quad x < -\frac{a}{2} \quad (\text{I})$$

$$V(x) = 0 \quad -\frac{a}{2} \leq x \leq \frac{a}{2} \quad (\text{II}) \quad (6-1)$$

$$V(x) = U \quad x > \frac{a}{2} \quad (\text{III})$$

We have selected this particular representation of the potential function since it is symmetric in  $x$ :

$$V(x) = V(-x) \quad (6-2)$$

The corresponding Hamiltonian operator  $H_{\text{op}}$  is also symmetric:

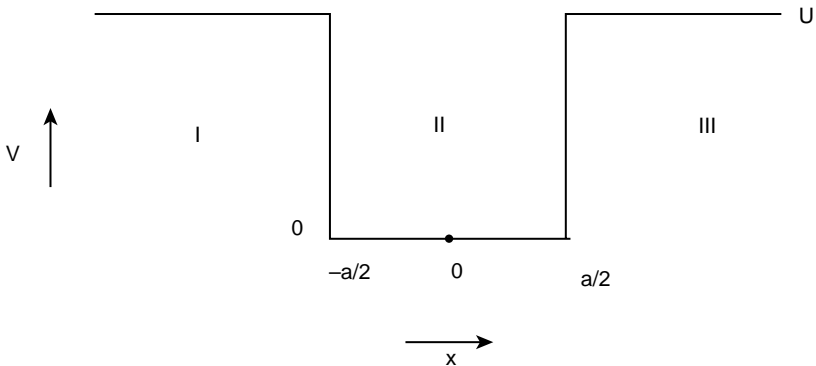
$$H_{\text{op}}(x) = H_{\text{op}}(-x) \quad (6-3)$$

It is now easily shown that the eigenfunctions  $\Psi_n(x)$  of the above Hamiltonian must be either symmetric or antisymmetric in  $x$ . The nondegenerate eigenfunctions are defined as

$$H_{\text{op}}(x)\Psi_n(x) = E_n\Psi_n(x) \quad (6-4)$$

It follows from Eq. (6-3) that we also have

$$H_{\text{op}}(x)\Psi_n(-x) = E_n\Psi_n(-x) \quad (6-5)$$



**Figure 6-1** Potential function of a particle in a one-dimensional finite box.

The two functions  $\psi_n(x)$  and  $\psi_n(-x)$  should be proportional to each other since the eigenstate is nondegenerate. We therefore have

$$\psi_n(x) == \sigma\psi_n(-x) \quad (6-6)$$

By substituting Eq. (6-6) into itself, we find that

$$\sigma^2 = 1 \quad (6-7)$$

or

$$\sigma = \pm 1 \quad (6-8)$$

We have shown that the Hamiltonian of Eq. (6-3) has two sets of eigenfunctions, namely, a set that is symmetric in  $x$  and a set that is antisymmetric in  $x$ . The above argument may also be extended to degenerate eigenstates.

The Schrödinger equation for the particle in a finite box is given by

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + U\psi = E\psi \quad (6-9)$$

in regions I and III and by

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi \quad (6-10)$$

in region II. We limit our discussion to bound states, which are characterized by the condition

$$0 < E < U \quad (6-11)$$

We introduce the following substitutions:

$$\begin{aligned} 2mE &= \mu^2\hbar^2 \\ 2m(U - E) &= \lambda^2\hbar^2 \end{aligned} \quad (6-12)$$

where  $\lambda$  and  $\mu$  are positive real quantities. The Schrödinger equations (6-9) and (6-10) then take the form

$$\begin{aligned} \frac{d^2\psi}{dx^2} &= \lambda^2\psi & (\text{I and III}) \\ \frac{d^2\psi}{dx^2} &= -\mu^2\psi & (\text{II}) \end{aligned} \quad (6-13)$$

The above differential equations were discussed in Section 2.II, and their general solutions are

$$\begin{aligned}\psi(x) &= A_I \exp(\lambda x) + B_I \exp(-\lambda x) & \text{(I)} \\ \psi(x) &= A_{II} \cos(\mu x) + B_{II} \sin(-\mu x) & \text{(II)} \\ \psi(x) &= A_{III} \exp(\lambda x) + B_{III} \exp(-\lambda x) & \text{(III)}\end{aligned}\tag{6-14}$$

These expressions may be simplified. We first note that

$$B_I = A_{III} = 0 \tag{6-15}$$

in order to satisfy the normalization condition. We now separate the eigenfunctions into two sets, the functions  $\psi_s(x)$  that are symmetric in  $x$  and the functions  $\psi_a(x)$  that are antisymmetric in  $x$ . In the symmetric case the expression (6-14) is reduced to

$$\begin{aligned}\psi_s(x) &= A \exp(\lambda x) & \text{(I)} \\ \psi_s(x) &= B \cos(\mu x) & \text{(II)} \\ \psi_s(x) &= A \exp(-\lambda x) & \text{(III)}\end{aligned}\tag{6-16}$$

and in the antisymmetric case it may be written as

$$\begin{aligned}\psi_a(x) &= C \exp(\lambda x) & \text{(I)} \\ \psi_a(x) &= D \sin(\mu x) & \text{(II)} \\ \psi_a(x) &= -C \exp(-\lambda x) & \text{(III)}\end{aligned}\tag{6-17}$$

The eigenvalues and eigenfunctions may now be derived from the conditions that both the functions and their derivatives must be continuous at the points  $x = a/2$  and  $x = -a/2$ . This leads to the following set of equations for the symmetric case:

$$\begin{aligned}A \exp(-a\lambda/2) &= B \cos(a\mu/2) \\ \lambda A \exp(-a\lambda/2) &= \mu B \sin(a\mu/2)\end{aligned}\tag{6-18}$$

and to a different set of equations for the antisymmetric case:

$$\begin{aligned}C \exp(-a\lambda/2) &= -D \sin(a\mu/2) \\ \lambda C \exp(-a\lambda/2) &= \mu D \cos(a\mu/2)\end{aligned}\tag{6-19}$$

The equations for the eigenvalues are obtained by dividing the two equations (6-18) for the symmetric case and by dividing the two equations (6-19) for the antisymmetric case. The symmetric case gives

$$\cot(a\mu/2) = (\mu/\lambda) \tag{6-20}$$

and the antisymmetric equation becomes

$$\operatorname{tg}(a\mu/2) = (-\mu/\lambda) \quad (6-21)$$

The two equations may be solved by means of two substitutions. The first substitution is

$$E = \rho^2 U \quad 0 \leq \rho \leq 1 \quad (6-22)$$

and the second substitution is

$$\rho = \sin \beta\pi \quad (6-23)$$

The two equations (6-20) and (6-21) may then be transformed to

$$\cos[(p\rho + \beta)\pi] = 0 \quad (6-24a)$$

and

$$\sin[(p\rho + \beta)\pi] = 0 \quad (6-24b)$$

where we have introduced the parameter

$$p^2 = 2m Ua^2/h^2 \quad (6-25)$$

The solutions of these equations are given by

$$p\rho + \beta = \beta + p \sin \beta = n + 1/2 \quad n = 0, 1, 2, \dots \quad (6-26a)$$

and by

$$p\rho + \beta = \beta + p \sin \beta = n \quad n = 1, 2, 3, \dots \quad (6-26b)$$

respectively. These equations may be solved by means of graphical methods but a detailed discussion of this procedure falls outside the scope of this book. The lowest eigenvalue belongs to a symmetric eigenfunction. There is at least one eigenvalue if

$$p\rho \geq \frac{1}{2} \quad (6-27)$$

The number of eigenvalues depends on the parameter  $p$ , and it will be finite if  $p$  is finite.

When the energy  $E$  is larger than  $U$ , the wave function  $\psi(x)$  in regions I and III has the general form

$$\psi(x) = A \exp(i\lambda x) + B \exp(-i\lambda x) \quad (6-28)$$

which represents the motion of a free particle. The first term corresponds to a wave moving from left to right, and the second term corresponds to a wave moving in the opposite direction. Every value of  $E$  is permitted since they all correspond to acceptable solutions of the Schrödinger equations. The specific form of the wave functions is, of course, affected by the change in potential in region II, and again it depends on the continuity conditions at the points  $x = -a/2$  and  $x = a/2$ . We do not feel it necessary to present a detailed discussion of this situation since its mathematical analysis is rather involved and it does not contribute much to our understanding of the subject. In addition, it bears some resemblance to the case we will discuss in the next section.

### III. TUNNELING

The classical description of the encounter between a moving particle and a potential barrier is quite straightforward. If the kinetic energy of the particle is smaller than the height of the potential barrier, then the particle will be reflected since it is unable to pass through the barrier. If, on the other hand, its energy is larger than the height of the potential barrier, its motion is not significantly affected by the barrier when it continues to move across.

The quantum mechanical description of the same encounter between a particle and a potential barrier leads to predictions that exhibit essential differences from the classical model. According to quantum mechanics, the particle may pass through a potential barrier even if its energy is smaller than the height of the barrier. This effect is popularly known as tunneling.

Initially the tunneling effect was considered nothing more than an unusual and novel quantum mechanical prediction that had little practical significance. Since then, the effect has found many applications in solid state physics, and it is a key feature in the design and construction of a variety of electronic devices that are found in computers and other new technology applications. It is therefore a phenomenon that is of general interest.

The tunneling effect may be illustrated by means of the rectangular potential barrier sketched in Figure 6-2. It is defined as

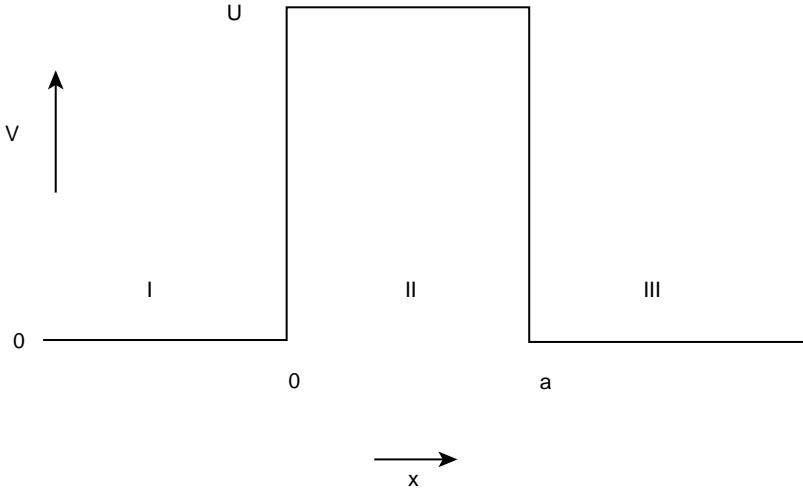
$$\begin{aligned} V(x) &= 0 & x < 0 & \quad \text{(I)} \\ V(x) &= U & 0 \leq x \leq a & \quad \text{(II)} \\ V(x) &= 0 & x > a & \quad \text{(III)} \end{aligned} \quad (6-29)$$

We shall see that there is no particular advantage in defining a potential that is symmetric in  $x$ .

We first consider the situation where the energy  $E$  of the particle is smaller than the height  $U$  of the potential barrier. The Schrödinger equation for region II may then be written as

$$\frac{d^2\psi}{dx^2} = \lambda^2\psi \quad \lambda^2 = \frac{2m(U - E)}{\hbar^2} \quad (6-30)$$





**Figure 6-2** The potential function used for the description of tunneling.

The Schrödinger equation for the other two regions, I and III, may be represented as

$$\frac{d^2\psi}{dx^2} = -\mu^2\psi \quad \mu^2 = \frac{2mE}{\hbar^2} \tag{6-31}$$

The solutions to these differential equations were presented in Section 2.II, and the general solution in region I is given by

$$\psi_I(x) = A_1 \exp(i\mu x) + B_1 \exp(-i\mu x) \tag{6-32}$$

The first term represents an incoming wave moving from left to right, and the second term represents a reflected wave moving from right to left. We therefore write the wave function in region I as

$$\psi_I(x) = \exp(i\mu x) + R \exp(-i\mu x) \tag{6-33}$$

where the coefficient  $R$  is related to the probability of reflection. By means of a similar argument, we write the wave function in region III as

$$\psi_{III}(x) = T \exp(i\mu x) \tag{6-34}$$

since in this region there is only one type of wave, the transmitted wave, moving from left to right. The coefficient  $T$  is related to the probability of transmission.

The solution of the wave function in region II is given by

$$\psi_{II}(x) = A \exp(\lambda x) + B \exp(-\lambda x) \tag{6-35}$$

The coefficients  $R$  and  $T$  may now be determined from the continuity conditions of the wave function and its derivative at the points  $x = 0$  and  $x = a$  by eliminating the coefficients  $A$  and  $B$ . At the point  $x = 0$  these conditions are

$$\begin{aligned} 1 + R &= A + B \\ i\mu(1 - R) &= \lambda(A - B) \end{aligned} \quad (6-36)$$

and at the point  $x = a$  they are

$$\begin{aligned} T \exp(i\mu a) &= A \exp(\lambda a) + B \exp(-\lambda a) \\ i\mu T \exp(i\mu a) &= \lambda A \exp(\lambda a) - B \exp(-\lambda a) \end{aligned} \quad (6-37)$$

We derive from the first set of equations (6-36) that

$$\begin{aligned} 2\lambda A &= (\lambda - i\mu)R + (\lambda + i\mu) \\ 2\lambda B &= (\lambda + i\mu)R + (\lambda - i\mu) \end{aligned} \quad (6-38)$$

and from the second set of equations (6-37) that

$$\begin{aligned} 2\lambda A &= (\lambda + i\mu) \exp(-\lambda a) T \exp(i\mu a) \\ 2\lambda B &= (\lambda - i\mu) \exp(\lambda a) T \exp(i\mu a) \end{aligned} \quad (6-39)$$

By setting Eqs. (6-38) and (6-39) equal to each other, we find that

$$\begin{aligned} (\lambda + i\mu) \exp(-\lambda a) T \exp(i\mu a) &= (\lambda - i\mu)R + (\lambda + i\mu) \\ (\lambda - i\mu) \exp(\lambda a) T \exp(i\mu a) &= (\lambda + i\mu)R + (\lambda - i\mu) \end{aligned} \quad (6-40)$$

We are primarily interested in the transmission coefficient  $T$ , which is given by

$$T \exp(i\mu a) = \frac{2i\lambda\mu}{(\mu^2 - \lambda^2) \sinh \lambda a + 2i\lambda\mu \cosh \lambda a} \quad (6-41)$$

The probability  $D$  of tunneling through the potential barrier is given by the product of  $T$  and its complex conjugate, or

$$D = T \cdot T^* = \frac{4\lambda^2\mu^2}{(\mu^2 + \lambda^2)^2 \sinh^2 \lambda a + 4\lambda^2\mu^2} \quad (6-42)$$

When the product  $\lambda a$  is much larger than unity, the tunneling probability is small and it may be approximated as

$$D = 16 \frac{E}{u} \left( 1 - \frac{E}{u} \right) \exp(-2\lambda a) \quad (6-43)$$

According to classical mechanics, the motion of a particle with energy  $E$  is not affected by the presence of a potential barrier as long as the energy  $E$  is larger than the height  $U$  of the barrier. It is interesting to note that this is not the case for the quantum mechanical description. The quantum mechanical results may again be derived in that case by solving the Schrödinger equation and imposing boundary conditions for the solution and its derivative. We do not present the details of this derivation, but the results show that for energy values  $E$  that are slightly larger than  $U$ , there is still a significant probability that the particle is reflected by the barrier. For larger values of  $E$ , the transmission probability approaches unity.

The above mathematical considerations illustrate why tunneling is considered a quantum mechanical effect.

#### IV. THE HARMONIC OSCILLATOR

The harmonic oscillator is a relatively simple system that offers a realistic model for the motion of a bound particle. It has therefore acquired a great deal of popularity among physicists, who have used it as the basis for a variety of theoretical models. It may be recalled that Planck used the harmonic oscillator for the illustration of the quantization concepts that he proposed in 1900.

The one-dimensional harmonic oscillator is a particle of mass  $m$  oscillating back and forth around an origin at the point  $x = 0$ . In the classical description of its motion, it is at any time subject to a force  $F$  that tends to move it back towards the origin and that is proportional to the distance  $x$  from the origin:

$$F = -kx \quad (6-44)$$

where  $k$  is known as the force constant. According to Section 3.III, the potential energy of the particle is then given by

$$V = \frac{1}{2}kx^2 \quad (6-45)$$

and its Hamiltonian function is

$$H = \frac{p^2}{2m} + \frac{kx^2}{2} \quad (6-46)$$

where  $p$  is its momentum.

We presented the classical description of the harmonic oscillator in Section 3.IV, and we found that its solution may be represented as

$$x(t) = A \sin(\omega t) \quad (6-47)$$

where  $A$  is the amplitude and  $\omega$  is the angular frequency of the oscillatory motion. The classical energy of the harmonic oscillator is given by

$$E = \frac{1}{2}m A^2 \omega^2 \quad (6-48)$$

It is a continuous function of both the angular frequency  $\omega$  and the amplitude  $A$ .

The quantum mechanical description of the harmonic oscillator is derived from the solution of the corresponding Schrödinger equation:

$$H_{\text{op}}\psi = E\psi \quad (6-49)$$

The Hamiltonian is given by Eq. (6-46) so that the Schrödinger equation has the form

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{kx^2}{2} \psi = E\psi \quad (6-50)$$

The solution of this differential equation requires a number of substitutions. The first one is

$$\varepsilon = \frac{2mE}{\hbar^2} \quad (6-51)$$

and it leads to the following equation:

$$\frac{d^2\psi}{dx^2} + \varepsilon\psi - \frac{kmx^2}{\hbar^2} \psi = 0 \quad (6-52)$$

After the second substitution

$$x = y\sqrt{\alpha} \quad \alpha^2 = \frac{\hbar^2}{km} \quad (6-53)$$

the harmonic oscillator equation assumes the simple form

$$\frac{d^2\psi}{dy^2} + (\alpha\varepsilon - y^2)\psi = 0 \quad (6-54)$$

It is possible to transform this differential equation into the differential equation (2-8), for which solutions are known; they are the Kummer functions discussed in Section 2.III. We first substitute

$$y^2 = t \quad (6-55)$$

into Eq. (6-54) and we obtain

$$t \frac{d^2 \psi}{dt^2} + \frac{1}{2} \frac{d\psi}{dt} + \frac{\alpha \varepsilon - t}{4} \psi = 0 \quad (6-56)$$

Next, we separate the asymptotic part of the solution by substituting

$$\psi(t) = w(t) \exp\left(-\frac{1}{2}t\right) \quad (6-57)$$

This leads to the desired new form of the differential equation

$$t \frac{d^2 w}{dt^2} + \left(\frac{1}{2} - t\right) \frac{dw}{dt} + \frac{\alpha \varepsilon - 1}{4} w = 0 \quad (6-58)$$

The above equation is identical to the differential equation (2-8), and its two solutions are given by Eqs. (2-16) and (2-17):

$$w_1(y) = {}_1F_1\left(\frac{1 - \alpha \varepsilon}{4}; \frac{1}{2}; y^2\right) \quad (6-59)$$

and

$$w_2(y) = y {}_1F_1\left(\frac{3 - \alpha \varepsilon}{4}; \frac{3}{2}; y^2\right) \quad (6-60)$$

It is easily seen that the first solution  $w_1(y)$  is symmetric in the variable  $y$ , while the second solution  $w_2(y)$  is antisymmetric in  $y$ .

The corresponding solutions  $\psi_1(y)$  and  $\psi_2(y)$  of the original Schrödinger equation (6-54) are

$$\psi_1(y) = {}_1F_1\left(\frac{1 - \alpha \varepsilon}{4}; \frac{1}{2}; y^2\right) \exp(-y^2/2) \quad (6-61)$$

and

$$\psi_2(y) = y {}_1F_1\left(\frac{3 - \alpha \varepsilon}{4}; \frac{3}{2}; y^2\right) \exp(-y^2/2) \quad (6-62)$$

In general these solutions do not represent acceptable wave functions because they are not normalizable since the Kummer functions behave asymptotically as  $\exp(y^2)$  for large values of the variable  $y$ . Acceptable wave functions are obtained only if they are reduced to finite polynomials. The latter condition is satisfied when the

variable  $a$  of the Kummer function  ${}_1F_1(a; c; x)$  is a negative integer, and this is the condition that determines the eigenvalues of the harmonic oscillator. We consider the two cases of the symmetric and antisymmetric eigenfunctions separately.

The symmetric eigenfunctions are represented by the solution  $\psi_1(y)$  of Eq. (6-61), and the corresponding eigenvalues are derived by imposing the condition

$$\alpha\varepsilon = 4n + 1 \quad (6-63)$$

We write the corresponding energy eigenvalues as  $E_n^s$ , and they are obtained by substituting Eqs. (6-51) and (6-53) into Eq. (6-63):

$$E_n^s = \left(2n + \frac{1}{2}\right)\hbar\omega \quad n = 0, 1, 2, \dots, \text{ etc.} \quad (6-64)$$

The corresponding symmetric eigenfunctions  $\psi_n^s(y)$  are given by

$$\psi_n^s(y) = {}_1F_1\left(-n; \frac{1}{2}; y^2\right) \exp(-y^2/2) \quad (6-65)$$

The other set of eigenvalues corresponding to the antisymmetric eigenfunctions are derived in a similar fashion from the antisymmetric solution  $\psi_2(y)$  of Eq. (6-62). Instead of Eq. (6-63) we now impose the condition

$$\alpha\varepsilon = 4n + 3 \quad (6-66)$$

and we find that the energy eigenvalues  $E_n^a$  are now given by

$$E_n^a = \left(2n + 1 + \frac{1}{2}\right)\hbar\omega \quad n = 0, 1, 2, 3, \dots, \text{ etc.} \quad (6-67)$$

The corresponding antisymmetric eigenfunctions are given by

$$\psi_n^a(y) = y {}_1F_1\left(-n; \frac{3}{2}; y^2\right) \exp(-y^2/2) \quad (6-68)$$

We may combine the two different sets of eigenvalues and eigenfunctions by writing the eigenvalues of the harmonic oscillator as

$$E_k = \left(k + \frac{1}{2}\right)\hbar\omega \quad k = 0, 1, 2 \quad (6-69)$$

The corresponding eigenfunctions are

$$\begin{aligned} \psi_k(y) &= \psi_n^s(y) & \text{if } k = 2n \\ &= \psi_n^a(y) & \text{if } k = 2n + 1 \end{aligned} \quad (6-70)$$

The specific form of the harmonic oscillator eigenfunctions may be derived from the definition (2-3) of the confluent hypergeometric function  ${}_1F_1(a; c; x)$ . The (unnormalized) eigenfunctions belonging to the lowest four eigenvalues are

$$\begin{aligned}\psi_0(y) &= \exp(-y^2/2) \\ \psi_1(y) &= y \exp(-y^2/2) \\ \psi_2(y) &= (2y^2 - 1) \exp(-y^2/2) \\ \psi_3(y) &= (2y^3 - 3y) \exp(-y^2/2)\end{aligned}\tag{6-71}$$

It may be interesting to compare the quantum mechanical and classical descriptions of the harmonic oscillator. In the quantum mechanical description, only specific discrete energy eigenvalues are allowed, and according to Eq. (6-69), the oscillator has a finite nonzero energy in its lowest eigenstate. This energy with magnitude

$$E_0 = \frac{1}{2} \hbar \omega\tag{6-72}$$

is known as the zero-point energy of the oscillator. The existence of this zero-point energy can be explained from the Heisenberg indeterminacy principle. In the classical model, the harmonic oscillator energy is a continuous function of the amplitude  $A$  and the angular frequency  $\omega$ , and it may assume any positive value. The classical energy may even be equal to zero, but in that case the particle is no longer moving. It follows that the permitted energy values are quite different in the quantum mechanical and classical models.

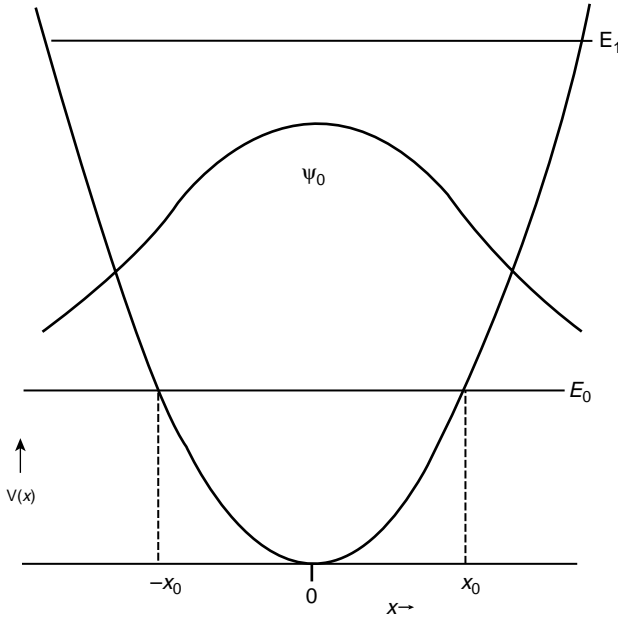
Another difference between the two models may be seen in Figure 6-3, where we have drawn the potential function  $V(x)$  of the harmonic oscillator as a function of the displacement coordinate  $x$  together with a horizontal line at the level of the lowest energy eigenvalue  $E_0$ . The points of intersection  $\pm x_0$  of the potential function and the line are given by

$$kx_0^2 = \hbar \omega\tag{6-73}$$

or

$$x_0 = \sqrt{\alpha}\tag{6-74}$$

We have also plotted the probability density of the particle belonging to the normalized eigenfunction  $\psi_0$  of Eq. (6-71) in Figure 6-3, and we see that the probability density is finite outside the range  $-x_0 \leq x \leq x_0$ . Within the range the energy is larger than the potential energy  $V(x)$  and the kinetic energy is positive, while outside the range  $E_0$  is smaller than  $V(x)$  and the kinetic energy is therefore negative. Negative kinetic energies are not allowed in classical mechanics, and in the



**Figure 6-3** The potential function  $V(x)$  and the lowest eigenvalue of the harmonic oscillator together with the corresponding eigenfunction.

classical description the motion of the particle is confined to oscillations between the points  $-x_0$  and  $x_0$ , which are the classical turning points of the oscillatory motion. The quantum mechanical model therefore allows the particle to move beyond the classical turning points, while the classical model does not.

We should finally mention that the matrices

$$x(n, m) = \langle \Psi_n(x) | x | \Psi_m | x \rangle \tag{6-75}$$

and

$$p(n, m) = -i\hbar \langle \Psi_n(x) | (d/dx) | \Psi_m | (x) \rangle \tag{6-76}$$

which played a prominent role in Heisenberg's matrix mechanics may be derived from the eigenfunctions (6-61) and (6-62) after normalizing the functions.

The results are

$$\begin{aligned} x(n, n - 1) &= (\alpha n/2)^{1/2} \\ x(n - 1, n) &= (\alpha n/2)^{1/2} \\ x(n, m) &= 0 \quad \text{if } m \neq n \pm 1 \end{aligned} \tag{6-77}$$



and

$$\begin{aligned}
 p(n, n-1) &= i\hbar(n/2\alpha)^{1/2} \\
 p(n-1, n) &= -i\hbar(n/2\alpha)^{1/2} \\
 p(n, m) &= 0 \quad \text{if } m \neq n \pm 1
 \end{aligned}
 \tag{6-78}$$

We do not present the derivations of the above expressions.

## V. PROBLEMS

- 6-1** Calculate the probability  $D$  for an electron with an energy of 0.5 eV to tunnel through a rectangular potential barrier with a height of 1 eV and a width of 1 Ångstrom unit.
- 6-2** Perform the same calculation as in Problem 6-1 for a proton instead of an electron.
- 6-3** Give analytical expressions for the normalized eigenfunctions of the four lowest eigenstates of the harmonic oscillator.
- 6-4** The symmetric eigenfunction  $\psi_n^s(y)$  of eq. (6-63) is zero for the values  $y = \pm y_1, \pm y_2, \dots$  etc. Calculate the sum of the squares of these zero values,  $\sum_i y_i^2$ .
- 6-5** Prove explicitly that the four lowest eigenfunctions of the harmonic oscillator are orthogonal.
- 6-6** Consider a particle of mass  $m$  moving in potential field that is defined as

$$V(x) = 0 \quad x < 0 \quad V(x) = U \quad x \geq 0$$

Assume that the particle is coming from the left with an energy  $E < U$ . Derive the wave function for this particle.

---

# 7

---

## ANGULAR MOMENTUM

### I. INTRODUCTION

The solution of the hydrogen atom Schrödinger equation presents a greater challenge than the problems we discussed before since it requires the solution of a three-dimensional differential equation rather than the one-dimensional equations of the previous chapter. The hydrogen atom is a special case of a particle moving in a central force field where the potential function  $V$  depends only on the distance  $r$  between the particle and the origin of the coordinate system.

We briefly discussed the classical description of a particle in a central force field in Section 3.V. We showed that it is convenient to introduce the angular momentum vector  $\mathbf{M}$  because  $\mathbf{M}$  is constant in time in the case of a central potential field. This feature makes it possible to separate the equations of motion in a radial and an angular part.

The classical equivalent of the hydrogen atom is the Kepler problem; in both cases, the potential function  $V(r)$  is inversely proportional to the coordinate  $r$ . Both the radial and angular equations of motion of the Kepler problem may be solved by making use of the facts that both the energy  $E$  and the angular momentum vector  $\mathbf{M}$  are constant in time.

By analogy, it is also possible to separate the Schrödinger equation of a particle in a central force field in a radial and an angular part by making use of the properties of the angular momentum. However, there are differences between the classical and quantum mechanical procedures. While in the classical theory we use the time

independence of the angular momentum, we use the commutation relations of the corresponding operators in the quantum theory. We discuss these in the following section.

In this chapter, we discuss the eigenvalues and eigenfunctions of the angular momentum operators. They are closely related to the angular part of the eigenfunctions of a particle in a central force field. They are also identical to the solutions of the rigid rotor, and it is therefore logical to discuss the latter system here. We postpone the solution of the radial part of the hydrogen atom Schrödinger equation to the next chapter.

## II. COMMUTING OPERATORS

The commutator of two operators  $F_{\text{op}}$  and  $G_{\text{op}}$  is defined as

$$[F_{\text{op}}, G_{\text{op}}] = F_{\text{op}} G_{\text{op}} - G_{\text{op}} F_{\text{op}} \quad (7-1)$$

If the commutator is zero, that is, if

$$F_{\text{op}} G_{\text{op}} = G_{\text{op}} F_{\text{op}} \quad (7-2)$$

the two operators are said to commute

Commuting operators are of special interest in quantum mechanics because they have common eigenfunctions. We define the eigenvalues and eigenfunctions of the two operators as

$$F_{\text{op}} f_n = \lambda_n f_n \quad G_{\text{op}} g_n = \mu_n g_n \quad (7-3)$$

And we have

$$F_{\text{op}} (G_{\text{op}} f_n) = G_{\text{op}} (F_{\text{op}} f_n) = \lambda_n (G_{\text{op}} f_n) \quad (7-4)$$

It follows therefore that the function  $G_{\text{op}} f_n$  is an eigenfunction of the operator  $F_{\text{op}}$  belonging to the eigenvalue  $\lambda_n$ . If the latter eigenvalue is nondegenerate, then the function  $G_{\text{op}} f_n$  must be proportional to  $f_n$ ; in other words

$$G_{\text{op}} f_n = \sigma f_n \quad (7-5)$$

It follows that  $f_n$  is also an eigenfunction of  $G_{\text{op}}$  if  $\lambda_n$  is a nondegenerate eigenvalue of  $F_{\text{op}}$ .

If the eigenvalue  $\lambda_n$  is degenerate, the argument becomes slightly more involved but the conclusion remains the same. The two operators  $F_{\text{op}}$  and  $G_{\text{op}}$  have common eigenfunctions if they commute.

### III. COMMUTATION RELATIONS OF THE ANGULAR MOMENTUM

It appears that the commutation relations between the components of the angular momentum defined in Section 3.V, and those between the angular momentum and the Hamiltonian operator, are helpful in deriving the quantum mechanical description of the system.

The classical expressions for the components of the angular momentum  $\mathbf{M}$  were described in Eq. (3-30). The corresponding quantum mechanical operators are then obtained by using the relations of Eq. (5-2):

$$\begin{aligned} M_x &= \frac{\hbar}{i} \left( y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \\ M_y &= \frac{\hbar}{i} \left( z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \\ M_z &= \frac{\hbar}{i} \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \end{aligned} \quad (7-6)$$

It is easily verified that each of the above three operator commutes with the Laplace operator  $\Delta$ :

$$[M_x, \Delta] = [M_y, \Delta] = [M_z, \Delta] = 0 \quad (7-7)$$

They also commute with a potential  $V(r)$  that depends only on the polar coordinate  $r$  of Section 3.VI:

$$[M_x, V(r)] = [M_y, V(r)] = [M_z, V(r)] = 0 \quad (7-8)$$

It follows therefore that each of the components also commutes with the Hamiltonian operator  $H_{\text{op}}$  of a particle in a central force field:

$$[M_x, H_{\text{op}}] = [M_y, H_{\text{op}}] = [M_z, H_{\text{op}}] = 0 \quad (7-9)$$

The magnitude of the angular momentum is represented by the operator  $M^2$ , which is defined as

$$M^2 = M_x^2 + M_y^2 + M_z^2 \quad (7-10)$$

It is easily shown from Eq. (7-9) that this operator also commutes with the Hamiltonian:

$$[M^2, H_{\text{op}}] = 0 \quad (7-11)$$

The commutation relations between the components of the angular momentum may be derived from their definitions by means of simple arithmetic; they are

$$\begin{aligned} [M_x, M_y] &= i\hbar M_z \\ [M_y, M_z] &= i\hbar M_x \\ [M_z, M_x] &= i\hbar M_y \end{aligned} \quad (7-12)$$

From these relations, it may also be derived that

$$[M^2, M_x] = [M^2, M_y] = [M^2, M_z] = 0 \quad (7-13)$$

It follows that the Hamiltonian operator  $H_{\text{op}}$  of a particle in a central force field, the operator  $M^2$ , and any one of the three components—for example,  $M_z$ —all commute. It may therefore be concluded from the general properties of commuting operators discussed in Section 7.II that the three operators  $H_{\text{op}}$ ,  $M^2$ , and  $M_z$  have common eigenfunctions. This feature will prove to be very helpful in solving the Schrödinger equation of a particle in a central force field.

#### IV. THE RIGID ROTOR

The simplest example of a particle moving in a three-dimensional central force field is a free particle whose motion is confined to the surface of a sphere with radius  $R$ . We shall see later that this system is mathematically equivalent to the rotational motion of a diatomic molecule; the latter is popularly known as the rigid rotor. Our system is represented by the Schrödinger equation

$$-\frac{\hbar}{2m}\Delta\psi = E\psi \quad (7-14)$$

with the additional restraint that the coordinate  $r$  must be equal to a constant value  $R$ . The eigenfunctions of this equation may be written as

$$\psi(x, y, z) = r^{-n}F_n(x, y, z) \quad (7-15)$$

where  $F_n(x, y, z)$  is an Euler polynomial of the  $n$ th degree in the Cartesian coordinates  $x$ ,  $y$ , and  $z$ . The variable  $r$  is, of course, the distance between the particle and the origin.

An Euler polynomial is defined as a special case of a homogeneous polynomial, and the latter is a linear combination of all possible products of the variables  $x$ ,  $y$ , and  $z$  subject to the condition that the sum of the exponents of each product is equal to a constant value  $n$ . We may write the definition of a homogeneous polynomial of the  $n$ th degree as

$$\begin{aligned} f_n(x, y, z) &= \sum_j \sum_k \sum_l a(j, k, l)x^j y^k z^l \\ j + k + l &= n \end{aligned} \quad (7-16)$$

In our subsequent argument we will make use of an interesting property of homogeneous polynomials. If we define the operator  $\Lambda$  as

$$\Lambda = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + z \frac{\partial}{\partial z} \quad (7-17)$$

then it is easily verified that

$$\Lambda f_n(x, y, z) = n f_n(x, y, z) \quad (7-18)$$

for every homogeneous polynomial of the  $n$ th degree. It should also be noted that the polynomial has

$$(n+1) + n + (n-1) + (n-2) + \cdots + 1 = (n+1)(n+2)/2 \quad (7-19)$$

different, linearly independent coefficients.

An Euler polynomial of the  $n$ th degree  $F_n(x, y, z)$  is now defined as a homogeneous polynomial  $f_n(x, y, z)$  that satisfies the condition

$$\Delta F_n(x, y, z) = 0 \quad (7-20)$$

It may be seen that  $\Delta F_n$  is a homogeneous polynomial of the  $(n-2)$ th degree. According to Eq. (7-19) it has  $n(n-1)/2$  coefficients. We require each of these coefficients to be zero, and we therefore have  $n(n-1)/2$  conditions for the coefficients of the homogeneous polynomial. It follows that the Euler polynomial  $F_n$  has

$$[(n+1)(n+2) - n(n-1)]/2 = 2n+1 \quad (7-21)$$

linearly independent coefficients.

Let us now show that the function  $\psi$  of Eq. (7-15) is an eigenfunction of the Schrödinger equation. We have

$$\Delta \psi = r^{-n} \cdot \Delta F_n + \left( \frac{2}{r} \frac{\partial}{\partial r} \frac{1}{r^n} \right) (\Lambda F_n) + F_n \Delta r^{-n} \quad (7-22)$$

By making use of Eqs. (7-18), (7-20), and (3-51), this equation may be reduced to

$$\begin{aligned} \Delta \psi &= F_n(x, y, z) [-2n^2 + n(n+1) - 2n] r^{-n-2} \\ &= -n(n+1) r^{-2} \psi \end{aligned} \quad (7-23)$$

Since the polar coordinate  $r$  must be equal to the constant value  $R$ , substitution into the Schrödinger equation (7-14) gives

$$\begin{aligned} \Delta [r^{-n} F_n(x, y, z)] &= -[(n(n+1)/R^2] [r^{-n} F_n(x, y, z)] \\ &= (-2mE/\hbar^2) [r^{-n} F_n(x, y, z)] \end{aligned} \quad (7-24)$$

It follows that the energy eigenvalues  $E_n$  are given by

$$E_n = n(n+1)(\hbar^2/2mR^2) \quad n = 0, 1, 2, 3, 4 \quad (7-25)$$

These eigenvalues are  $(2n+1)$ -fold degenerate, and the corresponding eigenfunctions are

$$\psi_n = r^{-n}F_n(x, y, z) \quad (7-26)$$

where  $F_n$  is an Euler polynomial of the  $n$ th degree. It is easily verified that the eigenfunctions  $\psi_n$  do not depend on the coordinate  $r$ ; they depend only on the polar angles  $\theta$  and  $\phi$  defined in Section 3.VI.

## V. EIGENFUNCTIONS OF THE ANGULAR MOMENTUM

The eigenfunctions of the angular momentum operator  $M^2$  are identical to the eigenfunctions derived in the previous section. The angular momentum of a particle moving on the surface of a sphere is given by

$$M = Rp \quad (7-27)$$

according to the definition (3-30) since the position vector  $\mathbf{r}$  is perpendicular to the momentum vector  $\mathbf{p}$ . The Hamiltonian operator  $H_{\text{op}}$  of a particle moving on a sphere is therefore given by

$$H_{\text{op}} = \frac{p^2}{2m} = \frac{(M^2)_{\text{op}}}{2mR^2} \quad (7-28)$$

Since

$$H_{\text{op}}[r^{-n}F_n(x, y, z)] = n(n+1)\frac{\hbar^2}{2mR^2}[r^{-n}F_n(x, y, z)] \quad (7-29)$$

according to Eqs. (7-14) and (7-24), it follows easily that

$$(M^2)_{\text{op}}[r^{-n}F_n(x, y, z)] = n(n+1)\hbar^2[r^{-n}F_n(x, y, z)] \quad (7-30)$$

It may be interesting to present an alternative derivation of the eigenvalues and eigenvectors of the angular momentum by making use of the polar coordinates introduced in Sections 3.V and 3.VI. The transformation of the angular momentum

operators (7-6) into polar coordinates gives

$$\begin{aligned} M_x &= i\hbar \left( \sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \\ M_y &= i\hbar \left( -\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \\ M_z &= -i\hbar \frac{\partial}{\partial \phi} \end{aligned} \quad (7-31)$$

By combining these results, we obtain

$$(M^2)_{\text{op}} = -\hbar^2 \left( \frac{\partial^2}{\partial \theta^2} + \frac{\cos \theta}{\sin \theta} \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) \quad (7-32)$$

A comparison of this result with Eq. (3-52) confirms once again the equivalence of this operator with the Hamiltonian of the rigid rotor.

Since the operators of the three vector components  $M_x$ ,  $M_y$ , and  $M_z$  all commute with the operator  $(M^2)_{\text{op}}$ , we may use this relation to further classify the eigenfunctions of  $(M^2)_{\text{op}}$ . The obvious choice of the three is  $M_z$  because of its mathematical simplicity.

The eigenvalue problem of the operator  $M_z$  is given by

$$(M_z)_{\text{op}} \chi = -i\hbar \frac{\partial \chi}{\partial \phi} = \lambda \chi \quad (7-33)$$

The solutions of this equation are

$$\chi(\phi) = \exp(i\lambda\phi/\hbar) \quad (7-34)$$

The eigenvalues are obtained by imposing the condition that the function  $\chi(\phi)$  is single-valued, namely,

$$\chi(\phi + 2\pi) = \chi(\phi) \quad (7-35)$$

This condition is satisfied if

$$(\lambda/\hbar) = m \quad m = 0, \pm 1, \pm 2, \dots \quad (7-36)$$

where  $m$  must be a positive or negative integer. The eigenvalues and eigenfunctions of the operator  $(M_z)_{\text{op}}$  are therefore given by

$$(M_z)_{\text{op}} \exp(im\phi) = m\hbar \exp(im\phi) \quad (7-37)$$



where  $m$  is an integer. The allowed values of the quantum number  $m$  are restricted to the range

$$m = 0, \pm 1, \pm 2, \dots \pm n \quad (7-38)$$

since the projection  $M_z$  cannot be larger than the magnitude  $M$  of the angular momentum vector.

In Eq. (7-21), we derived that the Euler polynomial  $F_n(x, y, z)$  has  $(2n + 1)$  linearly independent coefficients and that the eigenfunction  $\psi_n$  of the operator  $(M^2)_{\text{op}}$  therefore corresponds to an eigenvalue that is  $(2n + 1)$ -fold degenerate. By choosing appropriate parameters, we can construct a set of eigenfunctions  $\psi(n, m)$  that are eigenfunctions of both operators  $(M^2)_{\text{op}}$  and  $(M_z)_{\text{op}}$ . We list those eigenfunctions separately.

The Euler polynomial  $F_0$  is given by

$$F_0(x, y, z) = a \quad (7-39)$$

And it corresponds to the (unnormalized) eigenfunction

$$\psi(0, 0) = 1 \quad (7-40)$$

The Euler polynomial  $F_1$  is given by

$$F_1 = ax + by + cz \quad (7-41)$$

And it produces three eigenfunctions

$$\begin{aligned} \psi(1, 1) &= r^{-1}(x + iy) = \sin \theta \exp(i\phi) \\ \psi(1, 0) &= r^{-1}z = \cos \theta \\ \psi(1, -1) &= r^{-1}(x - iy) = \sin \theta \exp(-i\phi) \end{aligned} \quad (7-42)$$

The Euler polynomial  $F_2$  may be written as

$$F_2 = a(r^2 - 3z^2) + b(x^2 - y^2) + c_1xy + c_2xz + c_3z \quad (7-43)$$

By choosing appropriate parameters, we derive the following set of eigenfunctions:

$$\begin{aligned} \psi(2, 2) &= r^{-2}(x + iy)^2 = \sin^2 \theta \exp(2i\phi) \\ \psi(2, 1) &= r^{-2}z(x + iy) = \sin \theta \cos \theta \exp(i\phi) \\ \psi(2, 0) &= r^{-2}(r^2 - 3z^2) = 1 - 3 \cos^2 \theta \\ \psi(2, -1) &= r^{-2}z(x - iy) = \sin \theta \cos \theta \exp(-i\phi) \\ \psi_{2, -2} &= r^{-2}(x - iy)^2 = \sin^2 \theta \exp(-2i\phi) \end{aligned} \quad (7-44)$$

The eigenfunctions belonging to larger values of the angular momentum quantum number  $n$  may be derived in a similar fashion by first deriving the corresponding Euler polynomial and then combining the coefficients to obtain eigenfunctions of the operator  $(M_z)_{\text{op}}$ .

We should note that the operators corresponding to the vector components  $M_x$  and  $M_y$  also commute with  $(M^2)_{\text{op}}$ , and it is therefore possible to derive a set of functions that are eigenfunctions of either  $(M^2)_{\text{op}}$  and  $(M_x)_{\text{op}}$  or  $(M^2)_{\text{op}}$  and  $(M_y)_{\text{op}}$ . However, the corresponding mathematical derivations are much more complex than in the case of  $(M^2)_{\text{op}}$  and  $(M_z)_{\text{op}}$  discussed here, and they are beyond the scope of this book.

## VI. CONCLUDING REMARKS

Angular momentum has played an important role in the interpretation of atomic and molecular structure. Since the angular momentum operator commutes with the Hamiltonian of an atom, both operators have common eigenfunctions and each atomic eigenstate is characterized both by its energy eigenvalue and by the value of its angular momentum.

It was once customary to interpret atomic spectra by identifying the atomic eigenstates according to their angular momenta even before quantum mechanics was fully developed. This approach was known as the vector model of atomic structure, and it was widely used at the beginning of the twentieth century. We should also remind the reader that the proposals by Ehrenfest and Sommerfeld to quantize the angular momentum preceded the introduction of the Schrödinger equation. For these reasons, it is useful to have a good understanding of the quantum mechanical description of angular momentum.

We will also use the quantum mechanical results of the rigid rotor in Chapter 12 for the description of the nuclear motion in diatomic and polyatomic molecules.

## VII. PROBLEMS

**7-1** Prove explicitly the validity of the commutator relations

$$[M_x, M_y] = i\hbar M_z$$

$$[M_y, M_z] = i\hbar M_x$$

$$[M_z, M_x] = i\hbar M_y$$

**7-2** Prove explicitly that all three components  $M_x$ ,  $M_y$  and  $M_z$  commute with  $M^2$ .

**7-3** If we define the operators

$$M_1 = M_x + iM_y \quad M_{-1} = M_x - iM_y$$

prove that

$$[M_z, M_1] = \hbar M_1 \quad [M_z, M_{-1}] = -\hbar M_{-1}$$

**7-4** Prove also that

$$M_1 M_{-1} = M^2 - M_z^2 + \hbar M_z$$

$$M_{-1} M_1 = M^2 - M_z^2 - \hbar M_z$$

**7-5** If we denote the eigenfunctions of the angular momentum operators  $M^2$  and  $M_z$  by  $\psi(l, m)$  prove that

$$(m' - m - 1) \langle \psi(l, m') | M_1 | \psi(l, m) \rangle = 0$$

**7-6** Derive the three eigenfunctions of the operator  $M_x$  corresponding to the eigenvalue  $l = 1$  of  $M^2$ . These functions should be linear combinations of the functions  $\psi(1, 1)$ ,  $\psi(1, 0)$  and  $\psi(1, -1)$ .

**7-7** Give the general expression of the Euler polynomial  $F_3(x, y, z)$  of order 3. Derive from this expression the 7 (unnormalized) eigenfunctions of  $M^2$  that are also eigenfunctions of the operator  $M_z$ .

---

# 8

---

## THE HYDROGEN ATOM

### I. INTRODUCTION

As soon as Erwin Schrödinger derived his famous differential equation to describe the motion of bound particles according to wave mechanical principles, he set out to apply it to the hydrogen atom.

The hydrogen atom consists of two particles, a heavy proton with mass  $M$  and a much lighter electron with mass  $m$ . The motion of the two particles may be separated into two parts by means of a coordinate transformation. The first part is the motion of the center of gravity of the two particles. The center of gravity behaves as a free particle with mass  $(m + M)$ , and its motion is of no further interest to us. The second part is the relative motion of the two particles, which is described by a three-dimensional coordinate  $\mathbf{r}$  that represents the distance between the proton and the electron and by a corresponding conjugate momentum  $\mathbf{p}$  that is defined as

$$\mathbf{p} = \mu \mathbf{r} \quad \frac{1}{\mu} = \frac{1}{m} + \frac{1}{M} \quad (8-1)$$

Here  $\mu$  is known as the reduced mass of the two particles.

The Coulomb attraction between the proton and the electron is represented by a potential function

$$V = -\frac{e^2}{r} \quad (8-2)$$

and the Hamiltonian function of the hydrogen atom is therefore given by

$$H(\mathbf{r}, \mathbf{p}) = \frac{p^2}{2\mu} - \frac{e^2}{r} \quad (8-3)$$

The corresponding hydrogen atom Schrödinger equation is given by

$$-\frac{\hbar^2}{2\mu} \Delta \Psi - \frac{e^2}{r} \Psi = E \Psi \quad (8-4)$$

by analogy with Eq. (1-15a). The Laplace operator  $\Delta$  is defined in Eq. (1-17), and its transformation into polar coordinates is given by Eq. (3-51).

The solution of the three-dimensional differential equation is not easy, but we have discussed the various mathematical techniques that lead to its solution in previous chapters. We showed in Section 7.IV that the Hamiltonian (8-4) of a particle in a central force field commutes with the angular momentum operator  $(M^2)_{\text{op}}$  and that the two operators have common eigenfunctions. This allows us to separate the differential equation into an angular and a radial part. The radial differential equation may then be solved by reducing it to the differential equation for Kummer's function, which we discussed in Section 2.III. We present the details of the solution of the equation in the following section.

## II. SOLVING THE SCHRÖDINGER EQUATION

In order to solve the hydrogen atom Schrödinger equation (8-4) we introduce polar coordinates, and we note that the operator  $(M^2)_{\text{op}}$  depends only on the polar coordinates  $\theta$  and  $\phi$  and not on the variable  $r$ . We also note that the expression (7-32) of  $(M^2)_{\text{op}}$  in terms of the polar angle  $\theta$  and  $\phi$  is identical to the angular part of Eq. (3-52), where we expressed the Laplace operator in terms of the polar coordinates  $r$ ,  $\theta$ , and  $\phi$ . If we substitute these results into the Schrödinger equation (8-4), we obtain

$$-\frac{\hbar^2}{2\mu} \left( \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) \Psi + \frac{M^2}{2\mu r^2} \Psi - \frac{e^2}{r} \Psi = E \Psi \quad (8-5)$$

The eigenfunctions of the operator  $M^2$  may now be represented as

$$M^2 \psi(l, m; \theta, \phi) = l(l+1) \hbar^2 \psi(l, m; \theta, \phi) \quad (8-6)$$

Here we have introduced the customary symbol  $l$  to denote the eigenvalues of  $M^2$ . The eigenfunctions are defined in Section 7.IV. They are derived from Euler polynomials, and we have indicated that they depend on the polar angles  $\theta$  and  $\phi$  only, and not on the variable  $r$ .

It is now possible to separate the Schrödinger equation (8-5) into a radial and an angular part by representing the solution  $\Psi$  as

$$\Psi(r, \theta, \phi) = g_l(r) \Psi(l, m; \theta, \phi) \quad (8-7)$$

The angular part of the solution is an eigenfunction of both operators  $M^2$  and  $M_z$ , and the equation for the radial function  $g_l(r)$  becomes

$$-\frac{\hbar^2}{2\mu} \left[ \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{l(l+1)}{r^2} \right] g_l(r) - \frac{e^2 g_l(r)}{r} = E g_l(r) \quad (8-8)$$

In order to solve this equation, it is helpful to make the following substitutions:

$$E = -\frac{1}{2\sigma^2} \cdot \frac{e^2}{a_H} \quad r = \frac{\sigma a_H \rho}{2} \quad a_H = \frac{\hbar^2}{\mu e^2} \quad (8-9)$$

The quantity  $a_H$  is known as the Bohr radius. It has an approximate value of  $0.529 \times 10^{-8}$  cm, and it is a universal unit of length in calculations relating to atomic and molecular structure. The other substitutions of Eq. (8-9) are helpful because they eliminate some of the constants in the equation. They also have the effect of replacing the energy parameter  $E$  by a different parameter  $\sigma$ . Substitution of Eq. (8-9) into Eq. (8-8) reduces the equation to the following form:

$$\frac{d^2 g_l(\rho)}{d\rho^2} + \frac{2}{\rho} \frac{d g_l(\rho)}{d\rho} + \left[ -\frac{1}{4} + \frac{\sigma}{\rho} - \frac{l(l+1)}{\rho^2} \right] g_l(\rho) = 0 \quad (8-10)$$

The above equation may be further transformed into Kummer's differential equation (3-8) by means of two successive substitutions. The first one is

$$g_l(\rho) = \rho^l w_l(\rho) \quad (8-11)$$

which leads to the following equation:

$$\frac{d^2 w_l}{d\rho^2} + \frac{2l+2}{\rho} \frac{d w_l}{d\rho} + \left( -\frac{1}{4} + \frac{\sigma}{\rho} \right) w_l = 0 \quad (8-12)$$

The second substitution

$$w_l(\rho) = f_l(\rho) \exp(-\rho/2) \quad (8-13)$$

transforms the equation into the desired form

$$\rho \frac{d^2 f_l}{d\rho^2} + (2l+2-\rho) \frac{d f_l}{d\rho} - (l+1-\sigma) f_l = 0 \quad (8-14)$$

which is identical to Kummer's differential equation (2-8). Its solution is therefore

$$f_l(\rho) = {}_1F_1(l+1-\sigma; 2l+2; \rho) \quad (8-15)$$

The solution of the original hydrogen atom Schrödinger equation (8-10) is given by

$$g_l(\rho) = \rho^l \exp(-\rho/2) {}_1F_1(l+1-\sigma; 2l+2; \rho) \quad (8-16)$$

The second solution of Kummer's differential equation is not acceptable because it has a singularity at the origin  $\rho = 0$ .

It may be seen that the successive substitutions needed to solve the differential equation (8-8), and particularly the substitution (8-9), are far from obvious. It is therefore rumored that Schrödinger was helped by one of his mathematics colleagues at Zürich in solving the problem. This may or may not be true, but it should be noted that Schrödinger was quite capable of solving the problem by himself since he was a highly competent mathematician.

### III. DERIVING THE ENERGY EIGENVALUES

The energy eigenvalues of the hydrogen atom are derived from the solution of the corresponding Schrödinger equation by imposing the condition that this solution may be normalized—in other words, that it is finite everywhere. In accordance with this requirement, we have eliminated the second solution of the differential equation (8-14) since it has a singularity at the origin. The solution (8-16) is in general not acceptable because Kummer's function behaves asymptotically as  $\exp(\rho)$  for large values of  $\rho$  and it leads therefore to wave functions that are not normalizable. However, we have seen in Section 2.III that Kummer's function  ${}_1F_1(a; c; x)$  is reduced to a finite polynomial if the parameter  $a$  is equal to a negative integer. The solution (8-16) leads to acceptable wave functions only if the condition

$$l+1-\sigma = -v \quad v = 0, 1, 2, \dots, \text{ etc.} \quad (8-17)$$

is satisfied. This condition therefore determines the energy eigenvalues of the hydrogen atom.

Equation (8-17) is usually presented in a slightly different form:

$$\sigma = n \quad n \geq l+1 \quad (8-18)$$

which may also be formulated as

$$\begin{aligned} n &= 1, 2, 3, 4, \dots \\ l &= 0, 1, 2, \dots, n-1 \end{aligned} \quad (8-19)$$

In this manner, the stationary states of the hydrogen atom are characterized by two quantum numbers. The first quantum number  $n$  determines the energy, and the

second quantum number  $l$  determines the angular momentum corresponding to the stationary state. An additional third quantum number  $m$ , whose possible values are given by

$$m = 0, \pm 1, \pm 2, \dots \pm l \quad (8-20)$$

determines the eigenvalues of the operator  $M_z$  representing the projections of the angular momentum vector along the  $Z$  axis.

If we write the eigenfunctions of the hydrogen atom as  $\psi(n, l, m; r, \theta, \phi)$ , we may summarize these results as follows:

$$\begin{aligned} H_{\text{op}}\psi(n, l, m) &= -\frac{1}{2n^2} \cdot \frac{e^2}{a_H} \psi(n, l, m) \\ (M^2)_{\text{op}}\psi(n, l, m) &= l(l+1)\hbar^2\psi(n, l, m) \\ (M_z)_{\text{op}}\psi(n, l, m) &= m\hbar\psi(n, l, m) \end{aligned} \quad (8-21)$$

It should be noted that the Bohr radius  $a_0$  is actually defined as

$$a_0 = \frac{\hbar^2}{me^2} \quad (8-22)$$

where  $m$  is the electron mass. The corresponding atomic energy unit, the hartree, is defined as

$$\varepsilon_0 = \frac{e^2}{a_0} \quad (8-23)$$

In our calculation on the hydrogen atom, we use slightly different units of length and energy  $a_H$  and  $\varepsilon_H$  where the electron mass in  $m$  was replaced by the reduced mass  $\mu$  of Eq. (8-1).

**TABLE 8-1. Lower Energy Stationary States of the Hydrogen Atom and Their Names**

$n$	$l$	Name
1	0	1s
2	0	2s
2	1	2p
3	0	3s
3	1	3p
3	2	3d
4	0	4s
4	1	4p
4	2	4d
4	3	4f
etc.		



Each stationary state of the hydrogen atom is determined by the three quantum numbers  $n$ ,  $l$ , and  $m$ , but it is customary to describe them by means of a different notation that was developed in relation to the old spectroscopies theories. Here the values of the quantum number  $l$  are denoted by letters. The states  $l = 0$  are denoted by the letter  $s$ , the states  $l = 1$  by  $p$ ,  $l = 2$  by  $d$ ,  $l = 3$  by  $f$ ,  $l = 4$  by  $g$ , and so on. Each stationary state is then represented by a number corresponding to the value of  $n$  followed by a letter describing the value of  $l$ . We present the notation for the lower energy states of the hydrogen atom in Table 8-1. As may be seen, the quantum number  $m$  is usually not mentioned or indicated by a subscript.

#### IV. THE BEHAVIOR OF THE EIGENFUNCTIONS

The hydrogen atom eigenfunctions are obtained by combining the solution (8-16) of the radial part of the Schrödinger equation with the solution (7-30) of the angular part. We denote the eigenfunction by  $\psi_{n,l}$  and we find that

$$\psi_{n,l} = r^l \exp(-r/n) {}_1F_1(l+1-n; 2l+2; 2r/n) r^{-l} F_l(x, y, z) \quad (8-24)$$

where  $F_l(x, y, z)$  is an Euler polynomial of order  $l$ .

The specific form of the eigenfunctions may be obtained by substituting

$$\begin{aligned} l+1-n &= 0, -1, -2, \dots \text{etc.} \\ l &= n-1, n-2, \dots, 0 \end{aligned} \quad (8-25)$$

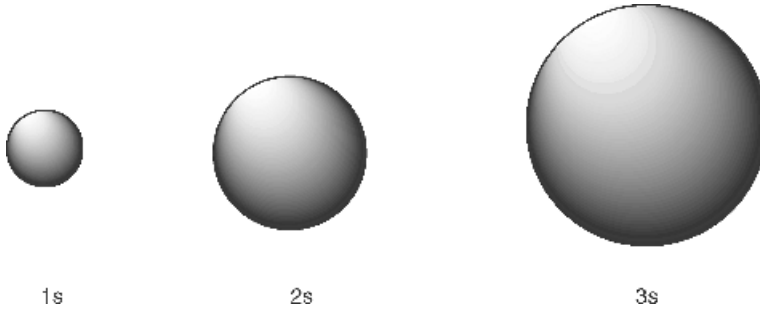
into the Kummer's function. The first few terms are given by

$$\begin{aligned} \psi_{n,n-1} &= \exp(-r/n) F_{n-1}(x, y, z) \\ \psi_{n,n-2} &= \left\{ 1 - \frac{r}{n(n-1)} \right\} \exp(-r/n) F_{n-2}(x, y, z) \\ \psi_{n,n-3} &= \left\{ 1 - \frac{2r}{n(n-2)} + \frac{2r^2}{n^2(n-2)(2n-3)} \right\} \exp(-r/n) F_{n-3}(x, y, z) \end{aligned} \quad (8-26)$$

We now list the detailed analytical expressions of the normalized eigenfunctions corresponding to the values  $n = 1, 2$ , and  $3$  of the principal quantum number  $n$ .

The eigenfunctions  $\psi(1s)$ ,  $\psi(2s)$ ,  $\psi(3s)$  do not depend on the polar angles  $\theta$  and  $\phi$  since the Euler polynomial  $F_0(x, y, z)$  is a constant. They are

$$\begin{aligned} \psi(1s) &= (1/\sqrt{\pi}) \exp(-r) \\ \psi(2s) &= (1/4\sqrt{2\pi})(r-2) \exp(-r/2) \\ \psi(3s) &= (1/81\sqrt{3\pi})(2r^2 - 18r + 27) \exp(-r/3) \end{aligned} \quad (8-27)$$



**Figure 8-1** Sketch of the probability density functions of the 1s, 2s, and 3s eigenstates of the hydrogen atom.

The probability densities corresponding to the  $ns$  states are spherically symmetric. It may be shown that the expectation values of the coordinate  $r$ , which are defined as

$$\langle r \rangle_{ns} = 4\pi \int \psi(ns)r \psi(ns)r^2 dr \quad (8-28)$$

have the value

$$\langle r \rangle_{ns} = (3n^2/2) \quad (8-29)$$

The probability densities of the  $ns$  states assume the form of spherically symmetric shells with a distance from the origin that is represented by Eq. (8-29). We have sketched the probability densities of the 1s, 2s, and 3s states in Figure 8-1.

The  $2p$  and  $3p$  eigenfunctions are both threefold degenerate since the Euler polynomial  $F_1(x, y, z)$  is given by

$$F_1(x, y, z) = ax + by + cz \quad (8-30)$$

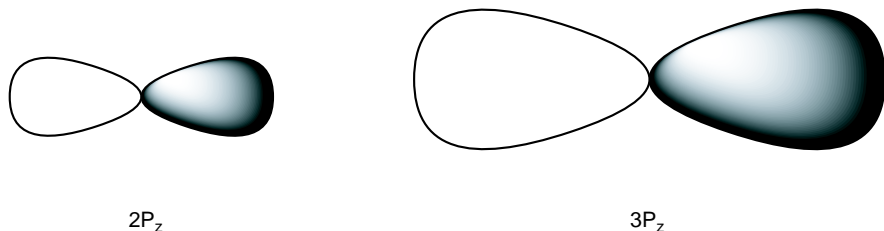
We denote the corresponding  $2p$  eigenfunctions by  $\psi(2p_x)$ ,  $\psi(2p_y)$ , and  $\psi(2p_z)$ , respectively. The analytical expressions of the normalized eigenfunctions are

$$\begin{aligned} \psi(2p_x) &= (1/4\sqrt{2\pi})x \exp(-r/2) \\ \psi(2p_y) &= (1/4\sqrt{2\pi})y \exp(-r/2) \\ \psi(2p_z) &= (1/4\sqrt{2\pi})z \exp(-r/2) \end{aligned} \quad (8-31)$$

The three degenerate  $3p$  eigenfunctions are defined in a similar fashion. The analytical expressions of the normalized functions are

$$\begin{aligned} \psi(3p_x) &= (\sqrt{2}/81\sqrt{\pi})(r-6)x \exp(-r/3) \\ \psi(3p_y) &= (\sqrt{2}/81\sqrt{\pi})(r-6)y \exp(-r/3) \\ \psi(3p_z) &= (\sqrt{2}/81\sqrt{\pi})(r-6)z \exp(-r/3) \end{aligned} \quad (8-32)$$

We present a sketch of the  $2p_z$  and  $3p_z$  eigenfunctions in Figure 8-2.



**Figure 8-2** Sketch of the hydrogen atom  $2p$  and  $3p$  eigenfunctions.

It should be noted that the above eigenfunctions of Eqs. (8-31) and (8-32) are not eigenfunctions of the operator  $M_z$ . However, they may be rearranged in different linear combinations in order to be transformed into a set of eigenfunctions of  $M_z$ , namely, as

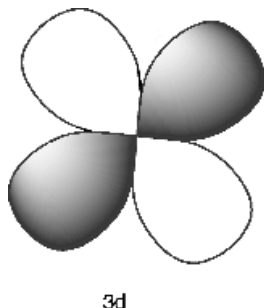
$$\begin{aligned}
 \psi(2p_1) &= (1/8\sqrt{\pi})r \sin \theta \exp(i\phi)\exp(-r/2) \\
 \psi(2p_0) &= (1/4\sqrt{2\pi})r \cos \theta \exp(-r/2) \\
 \psi(2p_{-1}) &= (1/8\sqrt{\pi})r \sin \theta \exp(i\phi)\exp(-r/2)
 \end{aligned} \tag{8-33}$$

The  $3p$  eigenfunctions may be transformed in a similar fashion. It should be noted that the set of  $M_z$  eigenfunctions are all complex functions. In practical computations of atomic and molecular structure there is a preference for real functions, and the real function of Eq. (8-31) that we sketched in Figure 8-2 rather than the complex functions of Eq. (8-33) are used as a basis for computational programs.

The  $3d$  eigenstate is fivefold degenerate since the Euler polynomial  $F_2(x, y, z)$  has five adjustable parameters. It is customary to represent the five normalized  $3d$  eigenfunctions as the following real expressions:

$$\begin{aligned}
 \psi(3d_{zz}) &= (1/8\sqrt{6\pi})(r^2 - 3z^2)\exp(-r/3) \\
 \psi(3d_{xz}) &= (\sqrt{2}/81\sqrt{\pi})xz \exp(-r/3) \\
 \psi(3d_{yz}) &= (\sqrt{2}/81\sqrt{\pi})yz \exp(-r/3) \\
 \psi(3d_{xy}) &= (\sqrt{2}/81\sqrt{\pi})xy \exp(-r/3) \\
 \psi(3d_{x^2-y^2}) &= (1/81\sqrt{2\pi})(x^2 - y^2)\exp(-r/3)
 \end{aligned} \tag{8-34}$$

In the case of the  $3d$  eigenfunctions, the angular distribution is much more interesting than the radial dependence. It is easily verified that  $\psi(3d_{xy})$  is zero in the planes  $x = 0$  and  $y = 0$  and that the planes  $x = \pm y$  contain the relative maxima and minima of the eigenfunction. The function  $\psi(3d_{x^2-y^2})$  exhibits the same behavior, and it may be derived from  $\psi(3d_{xy})$  by rotating the eigenfunction around the  $Z$  axis by  $45^\circ$ . In the latter case,  $\psi(3d_{x^2-y^2})$  is zero in the planes  $x = \pm y$ , and its maxima and minima are located along the  $X$  axis and the  $Y$  axis. Obviously,



**Figure 8-3** Sketch of the angular dependence of one of the  $3d$  eigenfunctions of the hydrogen atom.

the angular dependence of the functions  $\psi(3d_{xz})$  or  $\psi(3d_{yz})$  is similar to that of the function  $\psi(3d_{xy})$ . The function  $\psi(3d_{zz})$  is zero when  $3 \cos^2 \theta = 1$ , that is, on the surface of two cones making an angle of  $27.5^\circ$  with either the positive or negative  $Z$  axis from the origin. The relative maxima and minima of the latter eigenfunction are located along the  $Z$  axis and in the plane  $z = 0$ . The angular dependence of this eigenfunction is more complex than that of any of the other  $3d$  eigenfunctions. We have sketched the angular dependence of one of the  $3d$  eigenfunctions in Figure 8-3.

We will show that the hydrogen atom eigenfunctions play a major role in all atomic and molecular structure calculations and that their importance in present quantum chemical computations cannot be underestimated.

## V. PROBLEMS

- 8-1** The electric charge of an atomic nucleus is equal to  $Ze$  where  $Z$  is its atomic number, in the case of the hydrogen atom  $Z = 1$ . Derive the energy eigenvalues and the  $(1s)$  and  $(2s)$  eigenfunctions of a hydrogen like system where  $Z$  is different from unity.
- 8-2** Before the discovery of quantum mechanics the energy levels of the hydrogen atom were found to be  $E_n = R_H/n^2$  where  $R_H$  is the Rydberg constant of the hydrogen atom. Its value is

$$R_H = 109,677.58 \text{ cm}^{-1}$$

Calculate the lower excitation energy  $E(2p) - E(1s)$  of the hydrogen atom and of the  $\text{He}^+$  and  $\text{Li}^{2+}$  ions.

- 8-3** Prove that the hydrogen atom  $(3p_x)$  eigenfunction of Eq. (8-32) is normalized to unity.

- 8-4** Prove that the hydrogen atom ( $3s$ ) eigenfunction of Eq. (8-28) is normalized to unity.
- 8-5** Derive the detailed analytical expressions for the normalized ( $4s$ ) and ( $4p_z$ ) eigenfunctions of the hydrogen atom.
- 8-6** Calculate the expectation value of the distance  $r$  between the electron and the hydrogen nucleus for the ( $1s$ ), ( $2s$ ) and ( $2p_z$ ) eigenstates and compare the results.
- 8-7** Calculate the expectation values of  $(1/r)$ , the inverse of the distance between the electron and the hydrogen nucleus for the ( $1s$ ), ( $2s$ ) and ( $3s$ ) eigenstates and compare the results.
- 8-8** Calculate the expectation values of  $(1/r)$  for the ( $3s$ ), ( $3p_z$ ) and ( $3d_{zz}$ ) eigenstates of the hydrogen atom and compare the results.
- 8-9** Consider a particle of mass  $m$  in a three-dimensional spherical box where  $V(r) = 0$  if  $r \leq R$  and  $V(r) = \infty$  if  $r > R$ . Derive the energy eigenvalues and corresponding eigenfunctions for the case where the angular momentum of the particle is equal to zero.

---

# 9

---

## APPROXIMATE METHODS

### I. INTRODUCTION

In Schrödinger's series of papers where he first introduced his differential equation, he also presented the solution of his equation for the harmonic oscillator and for the hydrogen atom. Subsequently, an exact analytical solution of the Schrödinger equation of the hydrogen molecular ion was also presented in the literature. However, all efforts to derive analytical solutions of the Schrödinger equation for more complex systems, even for the helium atom, were unsuccessful.

There was, of course, a great deal of interest in applying quantum mechanics to more complex atomic and molecular systems, and this stimulated the development of approximate methods. Two different approximate procedures became the basis for all subsequent studies of atomic and molecular structure, namely, perturbation theory and the variational principle.

Perturbation theory is designed to predict the eigenvalues and eigenfunctions of a perturbed Hamiltonian  $H$  that differs only slightly from an unperturbed Hamiltonian  $H_0$  whose eigenvalues and eigenfunctions are known. The difference between  $H$  and  $H_0$  is called the perturbation.

We will show that the mathematical formalism of perturbation theory leads to exact predictions. On the other hand, its scope is rather limited since the only system for which the exact eigenvalues and eigenfunctions are known is that of the hydrogen atom. The ideal application of perturbation theory is a study of the effect of a homogeneous electric field on the eigenvalues and eigenfunctions of

the hydrogen atom. Perturbation theory was developed by Schrödinger in one of his original series of papers in order to calculate the hydrogen atom Stark effect, the interaction between an atom and a homogeneous electric field.

Perturbation theory has been widely used to calculate the effects of electric or magnetic fields on atoms and molecules. It is well suited for the derivation of formal expressions describing these effects. It should be realized, though, that in all these cases the unperturbed eigenvalues and eigenfunctions are not known, which makes the theoretical predictions subject to a great deal of uncertainty. Quantitative predictions of electric and magnetic properties are therefore less reliable than might be expected.

The variational method was initially derived from a mathematical theorem that formulates a relation between differential equations with boundary conditions, integral equations, and matrices of infinite order. In situations where exact solutions to the problem could not be found, the matrix representations led to convenient and fairly reliable approximate solutions. This latter procedure is now known as the variational method. It is particularly well suited for the use of computers. Nowadays almost all computer programs for calculations of atomic and molecular structure are based on the use of the variational principle.

Perturbation theory and the variational method may seem at first to be two different approximate procedures, but a more detailed analysis reveals that they are closely related. We will first present the variational principle, followed by perturbation theory. We discuss the relations between the two methods in the final section of this chapter.

## II. THE VARIATIONAL PRINCIPLE

We consider a differential equation with boundary conditions, and we define its eigenvalues and eigenfunctions by means of

$$H_{\text{op}} \psi_k = E_k \psi_k \quad k = 1, 2, 3, \dots \quad (9-1)$$

Here  $H_{\text{op}}$  is the operator defining the differential equation. The mathematical theorem mentioned in Section 9.I now states that any function  $f$  that satisfies the same boundary conditions as the eigenfunctions of Eq. (9-1) may be expanded in terms of those eigenfunctions:

$$f = \sum_k c_k \psi_k \quad (9-2)$$

We assume that the operator  $H_{\text{op}}$  is Hermitian. According to Eqs. (5-19) and (5-23), the eigenvalues  $E_k$  are then real and the eigenfunctions  $\psi_k$  are all orthogonal to one another. We also assume that the eigenfunctions  $\psi_k$  are normalized to unity. The coefficients  $c_k$  of Eq. (9-2) are then given by

$$c_k = \langle \psi_k | f \rangle \quad (9-3)$$

If any of the coefficients  $c_k$  in Eq. (9-2) are replaced by  $c_k + \delta c_k$ , then the function  $f$  changes by a small amount  $\delta f$ . It follows therefore that there is an infinite amount of possible variations  $\delta f$  of the function  $f$  since the expansion (9-2) contains an infinite number of coefficients  $c_k$ .

The variational principle now consists of the following statement: If  $g$  is a function satisfying the boundary conditions of Eq. (9-1) and if

$$\delta \left( \frac{\langle g | H | g \rangle}{\langle g | g \rangle} \right) = 0 \quad (9-4)$$

for all possible variations of the function  $g$ , then  $g$  is an eigenfunction of the operator  $H_{\text{op}}$ .

The variational principle is easily proved. We note that

$$\begin{aligned} \delta \langle g | H | g \rangle &= \langle \delta g | H | g \rangle + \langle g | H | \delta g \rangle \\ \delta \langle g | g \rangle &= \langle \delta g | g \rangle + \langle g | \delta g \rangle \end{aligned} \quad (9-5)$$

Substitution into Eq. (9-4) gives

$$\delta \frac{\langle g | H | g \rangle}{\langle g | g \rangle} = \frac{\langle \delta g | H - E | g \rangle + \langle g | H - E | \delta g \rangle}{\langle g | g \rangle} \quad (9-6)$$

with

$$E = \frac{\langle g | H | g \rangle}{\langle g | g \rangle} \quad (9-7)$$

Since

$$\langle g | H - E | \delta g \rangle = \langle \delta g | H - E | g \rangle^* \quad (9-8)$$

it follows immediately that

$$(H_{\text{op}} - E)g = 0 \quad (9-9)$$

In other words,  $g$  is an eigenfunction  $H_{\text{op}}$  with eigenvalue  $E$ .

In practical applications of the variational principle, it is helpful to make use of the following inequality. If  $f$  is an arbitrary function, then

$$\langle f | H | f \rangle \geq E_1 \langle f | f \rangle \quad (9-10)$$

where  $E_1$  is the smallest eigenvalue of the operator  $H_{\text{op}}$ . The equality (9-10) is again proved rather easily. We expand  $f$  in terms of the eigenfunctions  $\psi_k$  of  $H_{\text{op}}$



according to Eq. (9-1) and we find that

$$\begin{aligned} \langle f | H - E_1 | f \rangle &= \sum_m \sum_n c_m^* c_n \langle \psi_m | H - E_1 | \psi_n \rangle \\ &= \sum_m \sum_n c_m^* c_n (E_n - E_1) \langle \psi_m | \psi_n \rangle = \sum_n (E_n - E_1) c_n^* c_n \geq 0 \end{aligned} \quad (9-11)$$

The inequality (9-10) is quite useful for evaluating the accuracy of approximate eigenfunctions. For example, if a number of different approximate eigenfunctions are proposed for an atom or a molecule, then the eigenfunction that leads to the lowest energy expectation value is the most accurate of the group. Also, an approximate eigenfunction may depend on one or more unknown adjustable parameters. The best possible values of these parameters must be those values that correspond to the minimum of the energy expectation value.

### III. APPLICATIONS OF THE VARIATIONAL PRINCIPLE

Each atomic or molecular system is represented by a Hamiltonian operator, and the quantum mechanical description of the system may be derived either by solving the corresponding Schrödinger equation or by making use of the variational principle of the previous section. Neither of these approaches leads to exact solutions for any atom or molecule other than the hydrogen atom, but the variational method is better suited for the introduction of approximate procedures. Almost all computer programs used for the elucidation of atomic and molecular structure are therefore based on the variational principle rather than on the Schrödinger equation.

In order to apply the variational principle to the derivation of the eigenvalues and eigenfunctions of an atom or a molecule, it is necessary to define a known complete set of functions  $\phi_n$  that satisfies the same boundary conditions as the eigenfunctions  $\psi_k$  of the system. The term complete set means that any eigenfunction  $\psi$  of the system may be expanded in terms of the set of functions  $\phi_n$ :

$$\psi = \sum_{n=1}^{\infty} c_n \phi_n \quad (9-12)$$

According to Eqs. (9-4) and (9-6), the eigenfunctions and eigenvalues of the system are now obtained from the condition

$$\delta \langle \psi | H - E | \psi \rangle = 0 \quad (9-13)$$

rather than from the Schrödinger equation

$$H\psi = E\psi \quad (9-14)$$

which has (unknown) eigenvalues and eigenfunctions that we denote by  $\varepsilon_k$  and  $\psi_k$ :

$$H \psi_k = \varepsilon_k \psi_k \quad (9-15)$$

In order to solve the variational equation (9-13), we substitute the expansion (9-12) and we obtain

$$\delta \left\langle \sum_m \sum_n c_m^* c_n \langle \phi_m | H - E | \phi_n \rangle \right\rangle = 0 \quad (9-16)$$

We define the quantities  $H_{m,n}$  as

$$H_{m,n} = \langle \phi_m | H | \phi_n \rangle$$

and we rewrite the variational expression as

$$\delta < \sum_m \sum_n c_m^* c_n (H_{m,n} - E \delta_{m,n}) = 0 \quad (9-17)$$

for any possible variation in the function  $\psi$ . Here  $\delta_{m,n}$  are the Kronecker symbols we defined in Eq. (2-29). Since the expression (9-17) should apply to any possible variation in the function  $\psi$ , it should be valid for any variation in any of the coefficients  $c_m$ . This means that all derivatives of (9-17) with respect to these coefficients should be zero. For convenience, sake we differentiate with respect to  $c_m^*$  and we obtain the following infinite set of linear equations:

$$\sum_n (H_{m,n} - E \delta_{m,n}) c_n = 0 \quad m = 1, 2, 3, \dots \quad (9-18)$$

It is, of course, not possible to solve an infinite set of linear equations. However, it may be assumed that an approximate set of eigenvalues and eigenfunctions of the operator  $H$  may be derived by replacing the exact infinite expansion of Eq. (9-12) by a corresponding truncated expansion:

$$\psi(N) = \sum_{n=1}^N c_n(N) \phi_n \quad (9-19)$$

The infinite set (9-18) of linear equations then reduces to a finite set of  $N$  homogeneous linear equations with  $N$  unknowns:

$$\sum_{n=1}^N (H_{m,n} - E \delta_{m,n}) c_n(N) = 0 \quad m = 1, 2, \dots, N \quad (9-20)$$

This set of equations is identical to the equations (2-61) described in Section 2.VIII. These equations have a solution only if the determinant of the coefficients is equal to zero. The values  $E_k(N)$  of the parameter  $E$  for which the determinant is zero are defined as the eigenvalues of the matrix of the coefficients, and the corresponding solutions  $c_n(k, N)$  are the eigenvectors. The corresponding eigenfunctions  $\psi_k(N)$  are then defined as

$$\psi_k(N) = \sum_{n=1}^N c_n(k, N) \phi_n \quad (9-21)$$

According to the variational principle, the approximate eigenvalues  $E_k(N)$  and eigenfunctions  $\psi_k(N)$  will approach the exact eigenvalues  $\varepsilon_k$  and eigenfunctions  $\Psi_k$  when  $N$  tends toward infinity:

$$\begin{aligned} \lim_{N \rightarrow \infty} E_k(N) &= \varepsilon_k \\ \lim_{N \rightarrow \infty} \psi_k(N) &= \Psi_k \end{aligned} \quad (9-22)$$

It may also be proved that each approximate eigenvalue  $E_k(N)$  is always larger than or equal to the corresponding exact eigenvalue  $\varepsilon_k$  if both sets of eigenvalues are arranged in increasing magnitude:

$$E_1(N) \geq \varepsilon_1 \quad E_2(N) \geq \varepsilon_2 \quad E_3(N) \geq \varepsilon_3, \dots, \text{ etc.} \quad (9-23)$$

The computer hardware and software presently available are very efficient in evaluating the eigenvalues and eigenvectors of even large matrices. Most computer programs for calculating atomic and molecular structure are therefore based on the procedure described above. It is, of course, important to select appropriate functions as a basis for the wave function expansion in order to enhance its convergence. We will discuss the nature of the most suitable basis sets in subsequent chapters.

#### IV. PERTURBATION THEORY FOR A NONDEGENERATE STATE

The approximation method described in one of Schrödinger's early papers in order to derive the effect of a homogeneous electric field on the eigenvalues and eigenfunctions of the hydrogen atom is now known as Rayleigh-Schrödinger perturbation theory. Its purpose is the derivation of the eigenvalues and eigenfunctions of a Hamiltonian operator that may be represented as

$$H = H_0 + \lambda V \quad (9-24)$$

Here  $H$  and  $H_0$  are known as the perturbed and unperturbed Hamiltonian operators, respectively, and the difference  $\lambda V$  is called the perturbation. The latter contains a scaling parameter  $\lambda$  that is assumed to be small.

In the case analyzed by Schrödinger,  $H_0$  represented the hydrogen atom and its eigenvalues and eigenfunctions are known exactly. It should be noted that in most subsequent applications of perturbation theory to more complex atoms and molecules, the exact eigenvalues and eigenfunction of the unperturbed system are not known. Nevertheless, the formal derivation of perturbation theory is based on the assumption that the exact solutions of  $H_0$  are known. Quantitative results are then obtained by substituting approximate eigenvalues and eigenfunctions into the perturbation expressions.

In order to derive formal expressions for the various perturbation terms, we define the eigenfunctions and eigenvalues of the Hamiltonian operator  $H_0$  as

$$H_0 \psi_k = \varepsilon_k \psi_k \quad (9-25)$$

and those of the operator  $H$  as

$$H \phi_k = E_k \phi_k \quad (9-26)$$

We now consider the effect of the perturbation  $\lambda V$  on a nondegenerate eigenvalue of  $H_0$  that we denote as  $\varepsilon_0$  and on its eigenfunction  $\psi_0$ . The corresponding eigenvalue  $E_0$  and eigenfunction  $\phi_0$  of  $H$  may then be represented as power series in terms of the scaling parameter  $\lambda$ :

$$\begin{aligned} E_0 &= \varepsilon_0 + \lambda E'_0 + \lambda^2 E''_0 + \dots \\ \phi_0 &= \psi_0 + \lambda \phi'_0 + \lambda^2 \phi''_0 + \dots \end{aligned} \quad (9-27)$$

Substitution of these power series expansions into the Schrödinger equation (9-26) gives

$$(H_0 - \varepsilon_0 + \lambda V - \lambda E'_0 - \lambda^2 E''_0 \dots)(\psi_0 + \lambda \phi'_0 + \lambda^2 \phi''_0 + \dots) = 0 \quad (9-28)$$

The various perturbation equations are now derived by expanding Eq. (9-28) as a power series in terms of the scaling parameter  $\lambda$  and by setting each successive coefficient of this power series expansion equal to zero:

$$\begin{aligned} (H_0 - \varepsilon_0)\psi_0 &= 0 \\ (H_0 - \varepsilon_0)\phi'_0 + (V - E'_0)\psi_0 &= 0 \\ (H_0 - \varepsilon_0)\phi''_0 + (V - E'_0)\phi'_0 - E''_0\psi_0 &= 0, \text{ etc.} \end{aligned} \quad (9-29)$$

The first of the above perturbation equations is, of course, automatically satisfied. The other two equations may be simplified by multiplying them on the left by  $\psi_0$  and by subsequent integration:

$$\begin{aligned} \langle \psi_0 | H_0 - \varepsilon_0 | \phi'_0 \rangle + \langle \psi_0 | V - E'_0 | \psi_0 \rangle &= 0 \\ \langle \psi_0 | H_0 - \varepsilon_0 | \phi''_0 \rangle + \langle \psi_0 | V - E'_0 | \phi'_0 \rangle &= E''_0 \end{aligned} \quad (9-30)$$

Since  $H_0$  is Hermitian, the first terms are zero and we find that

$$\begin{aligned} E'_0 &= \langle \psi_0 | V | \psi_0 \rangle \\ E''_0 &= \langle \psi_0 | V - E'_0 | \phi'_0 \rangle \end{aligned} \quad (9-31)$$

The first-order energy perturbation  $E'_0$  is obtained by a straightforward integration, but the second-order term  $E''_0$  depends on the first-order wave function perturbation  $\phi'_0$ . The determination of the latter function requires solution of the inhomogeneous differential equation

$$(H_0 - \varepsilon_0)\phi'_0 = -(V - E'_0)\psi_0 \quad (9-32)$$

We discuss three different approaches to the solution of this differential equation.

We note first that the solution of an inhomogeneous differential equation is not unique since it is always permissible to add an arbitrary amount of the solution of the homogeneous equation (in this case  $\psi_0$ ) to any solution. However, we define a unique solution by imposing the condition

$$\langle \phi'_0 | \psi_0 \rangle = 0 \quad (9-33)$$

The best-known method for solving the perturbation equation (9-32) consists of expanding the unknown function  $\phi'_0$  in terms of the complete set of eigenfunctions  $\psi_k$  of the operator  $H_0$ :

$$\phi'_0 = \sum_{n=1}^{\infty} a_n \psi_n \quad (9-34)$$

We note that the expansion coefficient  $a_0$  is equal to zero because of the condition (9-33).

Substitution of the expansion (9-34) into the differential equation (9-32) gives

$$\sum_n a_n (H_0 - \varepsilon_0)\psi_n = -(V - E'_0)\psi_0 \quad (9-35)$$

The coefficients  $a_n$  are then obtained by multiplying the equation by one of the eigenfunctions  $\psi_k$  and subsequent integration:

$$\sum_n a_n (\varepsilon_n - \varepsilon_0) \langle \psi_k | \psi_n \rangle = -\langle \psi_k | V | \psi_0 \rangle \quad (9-36)$$

or

$$a_k = -\frac{\langle \psi_k | V | \psi_0 \rangle}{\varepsilon_k - \varepsilon_0} \quad (9-37)$$

The second-order energy perturbation  $E_0''$  is then derived by substituting this result into Eq. (9-31):

$$E_0'' = - \sum_k \frac{\langle \psi_0 | V | \psi_k \rangle \langle \psi_k | V | \psi_0 \rangle}{\epsilon_k - \epsilon_0} \quad (9-38)$$

It may be shown that the third-order energy perturbation may also be derived from the first-order wave function perturbation  $\phi_0'$  and that the  $(2n + 1)$ -th energy perturbation may be obtained by simple integration from the  $n$ th eigenfunction perturbation. The higher-order perturbation terms have become of interest lately because of experimental advances in nonlinear optics, but their derivation is beyond the scope of this book.

The perturbation expression (9-38) was used extensively as a basis for the discussion of electric and magnetic properties of molecules, but it offers only a qualitative representation of the various effects and it is not well suited for quantitative evaluations. Even in cases where approximate ground state wave functions are available, there is much less information about the excited state eigenfunctions, so numerical evaluations of the perturbation expression (9-38) present awkward problems.

As an interesting alternative, we will show how the perturbation equation (9-32) may also be solved by making use of the variational theorem of Eq. (9-10). We first note that by substituting Eq. (9-32) into Eq. (9-31), we may also write  $E_0''$  as

$$E_0'' = - \langle \phi_0' | H_0 - \epsilon_0 | \phi_0' \rangle \quad (9-39)$$

In the special case where  $\epsilon_0$  is the lowest eigenvalue of  $H_0$  we have

$$\langle g - \phi_0' | H - \epsilon_0 | g - \phi_0' \rangle \geq 0 \quad (9-40)$$

for any function  $g$ . We may rewrite Eq. (9-41) as

$$\langle g | H_0 - \epsilon_0 | g \rangle + \langle g | V - E_0' | \psi_0 \rangle + \langle \psi_0 | V - E_0' | g \rangle \geq E_0'' \quad (9-41)$$

by making use of Eqs. (9-32) and (9-39). This inequality presents a variational approach to the derivation of the second-order energy perturbation, especially in situations where we can make an educated guess about the nature of the perturbation function  $\phi_0'$ .

The third approach to the perturbation problem is to solve the differential equation analytically. This is, of course, the best approach, but unfortunately it is feasible in only a few cases. One of these is the perturbation of the hydrogen atom by a homogeneous electric field. We discuss this problem in the next section.

## V. THE STARK EFFECT OF THE HYDROGEN ATOM

At the beginning of the twentieth century, it was generally believed that the effect of a homogeneous electric field on atomic spectral lines was too small to be

measurable, but in 1913 *Johannes Stark* (1874–1957) observed both a splitting and a displacement in some spectral lines emitted by the hydrogen atom. The changes in atomic emission spectral lines became known as the Stark effect, and they were widely studied in subsequent years. It was found that the energy change in the lowest eigenstate of the hydrogen atom due to a homogeneous electric field  $F$  is proportional to the square of the electric field. This shift became known as the quadratic Stark effect as opposed to the more common linear Stark effect that was observed in many other cases.

The Schrödinger equation of a hydrogen atom in a homogeneous electric field is given by

$$-\frac{\hbar^2}{2\mu}\Delta\phi - \frac{e^2}{r}\phi - eFz\phi = E\phi \quad (9-42)$$

After the introduction of atomic units of length  $a_H$  and energy  $e^2/a_H$ , this equation is reduced to

$$-\frac{1}{2}\Delta\phi - \frac{1}{r}\phi - \lambda z\phi = E\phi \quad (9-43)$$

where the scaling parameter  $\lambda$  is given by

$$\lambda = a_H^2 F/e \quad (9-44)$$

The Stark effect of the hydrogen atom ground state corresponds obviously to the second-order energy perturbation, which is derived from the first-order perturbation equation (9-29)

$$\begin{aligned} (H_0 - \varepsilon_0)\phi'_0 + (V - E'_0)\psi_0 &= 0 \\ H_0 &= -\frac{1}{2}\Delta - \frac{1}{r} & V &= -z \\ \psi_0 &= (1/\sqrt{\pi})\exp(-r) & \varepsilon_0 &= -0.5 \end{aligned} \quad (9-45)$$

The above equation may be solved exactly by substituting

$$\phi'_0 = g \cdot \psi_0 \quad (9-46)$$

We have

$$\begin{aligned} \Delta(g \cdot \psi_0) &= \psi_0 \cdot (\Delta g) + g \cdot (\Delta \psi_0) \\ &+ 2\left(\frac{\partial \psi_0}{\partial x} \frac{\partial g}{\partial x} + \frac{\partial \psi_0}{\partial y} \frac{\partial g}{\partial y} + \frac{\partial \psi_0}{\partial z} \frac{\partial g}{\partial z}\right) \end{aligned} \quad (9-47)$$

Since

$$(H_0 - \varepsilon_0)\psi_0 = 0 \quad E'_0 = 0 \quad (9-48)$$

the perturbation equation (9-45) reduces to

$$\Delta g - 2\Omega g = -2z \quad (9-49)$$

$$\Omega = \frac{x}{r} \frac{\partial}{\partial x} + \frac{y}{r} \frac{\partial}{\partial y} + \frac{z}{r} \frac{\partial}{\partial z}$$

after dividing the equation by  $\psi_0$ .

In order to solve Eq. (9-49) we note that

$$\Delta(z) = 0 \quad \Delta(zr) = \frac{4z}{r} \quad (9-50)$$

$$\Omega(z) = \frac{z}{r} \quad \Omega(zr) = 2z$$

It follows that the function  $g$  may be represented as

$$g = az + b zr \quad (9-51)$$

and substitution into Eq. (9-49) gives

$$\frac{4zb}{r} - \frac{2az}{r} - 4bz = -2z \quad (9-52)$$

or

$$a = 1 \quad b = \frac{1}{2} \quad g = z + \frac{zr}{2} \quad (9-53)$$

The second-order energy perturbation is obtained as

$$E''_0 = \frac{1}{\pi} \iiint \left( z + \frac{zr}{2} \right) z \exp(-2r) dx dy dz = -\frac{9}{4} \quad (9-54)$$

This exact perturbation result is, of course, in perfect agreement with the experimental information derived from the Stark effect.

At first sight, this may seem an interesting result that is of little practical use since the hydrogen atom is the only system for which the perturbation equation may be solved directly. However, the first-order wave function perturbation described by Eq. (9-53) may be used as a basis for devising variational perturbation functions in order to solve molecular perturbation equations by means of the procedure described by Eq. (9-41). In fact, this approach has been used for the calculation of molecular polarizabilities.



## VI. PERTURBATION THEORY FOR DEGENERATE STATES

Many atomic spectral lines consist of multiplets. The multiplet structure may then be explained as a spectral transition between two groups of near-degenerate energy levels. In the theory of atomic structure, it was assumed that the two energy levels were initially degenerate but that they were subject to perturbations due to small interactions between orbital and spin angular momenta. Various empirical rules were proposed to explain the nature of these splittings. Eventually these rules were confirmed by applying the results of perturbation theory of degenerate states.

By analogy with Section 9.IV, we again consider the perturbation of an energy eigenvalue  $\varepsilon_0$ , but now there are  $N$  corresponding eigenfunctions  $\psi_{0,1}, \psi_{0,2} \dots \psi_{0,N}$  belonging to this eigenvalue rather than just one. We therefore have

$$(H_0 - \varepsilon_0)\psi_{0,k} = 0 \quad (9-55)$$

but also

$$(H_0 - \varepsilon_0) \sum_{k=1}^N a_k \psi_{0,k} = 0 \quad (9-56)$$

since any linear combination of the eigenfunctions  $\psi_{0,k}$  is also an eigenfunction.

If we now introduce a perturbation  $\lambda V$ , then the Schrödinger equation becomes

$$(H_0 - \varepsilon_0 + \lambda V - \lambda E'_0 - \lambda^2 E''_0 \dots) \left( \sum_k a_k \psi_{0,k} + \lambda \phi'_0 + \dots \right) = 0 \quad (9-57)$$

and the first-order perturbation equation takes the form

$$(H_0 - \varepsilon_0)\phi'_0 + (V - E'_0) \sum_k a_k \psi_{0,k} = 0 \quad (9-58)$$

In order to solve this equation, we multiply on the left by one of the eigenfunctions  $\psi_{0,m}^*$  and integrate. This yields the following set of linear equations:

$$\sum_{k=1}^N [V(m,k) - E'_0 \delta_{m,k}] a_k = 0 \quad m = 1, 2, \dots, N \quad (9-59)$$

$$V(m,k) = \langle \psi_{0,m} | V | \psi_{0,k} \rangle$$

We recognize that this is a set of  $N$  homogeneous linear equations with  $N$  variables—in other words, an eigenvalue problem of order  $N$ . The set of eigenvalues represents the first-order energy perturbations to the eigenvalue  $\varepsilon_0$ .

It should be noted that we could have obtained the same result by making use of the variational method, namely, by introducing a variational function

$$\Psi = \sum_k a_k \psi_{0,k} \quad (9-60)$$

to derive the eigenvalues of the Hamiltonian operator

$$H = H_0 + \lambda V \quad (9-61)$$

In this way, we would also determine the group of eigenvalues that are close to the eigenvalue  $\varepsilon_0$ . The advantage of using the variational approach is that it can also be applied to near-degenerate eigenvalues rather than eigenvalues that are exactly degenerate. Here we define near-degenerate eigenvalues as a group of eigenvalues that differ by amounts of the order  $\lambda$ , the scaling parameter of the perturbation.

In our experience, the variational approach is better suited to deal with perturbations of degenerate or near-degenerate states than a formal application of perturbation methods. Therefore, we will not discuss the perturbation theory of degenerate states in further detail.

## VII. CONCLUDING REMARKS

It is now possible to make fairly accurate predictions about the energy, geometry, and electronic structure of fairly large molecules with 50 or more atoms by making use of the various molecular structure computer programs. The best known of these are the Gaussian Program Packages, which are updated and improved almost every two years. These programs incorporate a great deal of work to develop and improve approximate methods in quantum mechanics used by many scientists over more than half a century. In this chapter, we have presented only a broad outline of the most basic principles underlying those efforts.

In subsequent chapters we will expand our discussion of approximate methods by presenting an approach that was specifically designed for dealing with many-electron systems namely the Hartree-Fock method, also known as Self Consistent Field Method. The derivation of this method is dependent on the exclusion principle and on the concept of electron spin which we have not yet discussed. It should also be noted that density functional methods have become quite popular in recent years. However a discussion of these and other more advanced and complex approximate methods falls outside the scope of this book.

## VIII. PROBLEMS

**9-1** If  $\varepsilon_1$  is the lowest and  $\varepsilon_2$  is the second lowest eigenvalue of a Hermitian operator  $H$  prove that

$$\langle \phi | H^2 | \phi \rangle \geq (\varepsilon_1 + \varepsilon_2) \langle \phi | H | \phi \rangle - \varepsilon_1 \varepsilon_2 \langle \phi | \phi \rangle$$

for any function  $\phi$ .

**9-2** The lowest eigenvalue of the hydrogen atom is  $\varepsilon_1 = -0.5$  and the corresponding (unnormalized) eigenfunction is  $\phi_1 = \exp(-r)$ . Determine the best possible value of the ground state energy of the hydrogen atom that can be

derived by minimizing the energy expectation value of the variational function

$$\psi = \exp(-ar^2)$$

with respect to the parameter  $a$ .

- 9-3** Prove that the second-order energy perturbation  $E_o''$  of the lowest eigenvalue is always positive.
- 9-4** We denote the third-order energy perturbation of a stationary state due to a perturbation  $\lambda V$  by  $E_o^{(3)}$ , consistent with Eq. (9-28). Prove that this energy perturbation may be expressed in terms of the first-order wave function perturbation as follows

$$E_o^{(3)} = \langle \phi_o' | V - E_o' | \phi_o' \rangle$$

- 9-5** Express the third-order energy perturbation  $E_o^{(3)}$  in terms of the unperturbed eigenvalues and eigenfunctions.
- 9-6** Prove that the fourth order energy perturbation  $E_o^{(4)}$  of the stationary state  $o$  due to a perturbation  $\lambda V$  may be written as

$$E_o^{(4)} = -\langle \phi_o'' | H_o - \epsilon_o | \phi_o'' \rangle - E_o'' \langle \phi_o' | \phi_o' \rangle$$

where the various quantities are defined in Eq. (9-28).

- 9-7** Calculate the second-order energy perturbation of the ground state of the hydrogen atom due to a homogeneous electric field along the Z axis by means of the variation-perturbation method described by Eq. (9-42) and by taking the variation function  $g$  as
- a)**  $g = az \exp(-r)$   
**b)**  $g = az \exp(-r/2)$   
 Compare the results.

- 9-8** Evaluate the first- and second-order energy perturbations of the ground state of a harmonic oscillator due to a perturbation  $V = x^3$ . The Hamiltonian of the harmonic oscillator is

$$H = (p^2/2m) + (kx^2/2).$$

- 9-9** Evaluate the first- and second order energy perturbations of the harmonic oscillator of problem 9-8 due to a perturbation  $V = x^4$ .

---

# 10

---

## THE HELIUM ATOM

### I. INTRODUCTION

The Schrödinger equation may easily be extended to make it applicable to many-electron systems. However, we have considered only one-electron systems up to this point because the wave functions of many-electron systems are subject to additional restraints. The most important of these is the Pauli exclusion principle that we briefly alluded to in Section 1.VI. It is equally important to include the electron spin in the quantum mechanical description of many-electron systems.

Both the exclusion principle and the existence of the electron spin were initially formulated in order to explain previous experimental discoveries. We mentioned in Section 1.VI that the exclusion principle was proposed by Pauli in 1925 and that the electron spin was introduced by Goudsmit and Uhlenbeck in the same year. It may therefore be helpful to give a brief description of the various experimental developments that preceded those theoretical proposals.

The exclusion principle and the electron spin apply to all many-electron systems, but in this chapter we only discuss the applications to two-electron systems and the helium atom in particular. Applications to more complex atoms and molecules are discussed in subsequent chapters.

## II. EXPERIMENTAL DEVELOPMENTS

We described in Section 1.III how the experimental information on the hydrogen atom spectrum helped Bohr to formulate the old quantum theory. We might add that the successful prediction of the hydrogen atom energy levels by Schrödinger offered solid support for the validity of his differential equation. At that time there was, of course, a great deal of experimental information available on the spectra of larger atoms and molecules. Of the many experimental discoveries at the beginning of the twentieth century, some were particularly relevant to the further development of quantum mechanics. In our opinion, the most important of these are (1) the discovery of X rays in 1895, (2) the discovery of the Zeeman effect in 1896, (3) the work of Stern and Gerlach on the splitting of molecular beams due to magnetic fields, and (4) the doublet structure of the spectra of alkali atoms. We discuss each of these separately.

The first Nobel Prize in physics was awarded in 1901 to *Wilhelm Conrad Röntgen* (1845–1923) for his discovery in 1895 of X rays, also known as Röntgen rays. We briefly described X rays in Section 1.VI since their experimental properties helped Louis de Broglie in his formulation of wave mechanics. We mentioned that X rays are electromagnetic waves with very short wavelengths of the order of 1 Å. Since X rays have such short wavelengths, their quanta have very high frequencies and consequently very high energies. It may therefore be assumed that the emission of an X ray quantum involves one of the inner electrons of an atom since only the electrons close to the atomic nucleus have energies that are comparable to the high energies of X rays. The frequencies of the X rays that were emitted by atoms were the major source of information on the energies of the inner electrons of various atoms.

Röntgen was a careful and meticulous experimentalist who built his own instruments and equipment. Much of the available information on the properties of X rays was derived from his accurate experiments during the decade after his initial discovery.

It was found that the X ray emission spectrum of a particular atom consists of various groups of lines. The group of lines with the highest frequency, corresponding to the energies of the most tightly bound electrons, was called the K-emission spectrum. The next group of lines is the L-emission spectrum. Then comes the M-emission spectrum, and so on. The group of electrons that give rise to the K lines is called the K shell. The next group is the L shell, and so on. It follows that the electrons do not all have the same energies. Instead they are divided into shells with quite different energies. This conclusion played an important role in the formulation of the exclusion principle.

An important discovery was made by *Henry Gwyn Jeffreys Moseley* (1887–1915), who had joined Rutherford at Cambridge as a graduate student. Moseley measured the Röntgen K lines of a large number of atoms, and he found that the frequencies of these lines could all be represented by the equation

$$\nu_K = T(Z - p)^2 \quad (10-1)$$

where  $Z$  is the atomic number or electric charge of the atomic nucleus and  $T$  and  $p$  are constants. Moseley's equation was subsequently used to determine the atomic number and the location in the periodic system of newly discovered elements. Moseley joined the British army during the First World War and was killed in 1915 on the Gallipoli Peninsula. He was considered one of the most promising young physicists in Europe, and his death was a tragic loss for science.

*Pieter Zeeman* (1856–1943) believed that an atomic spectral line might be affected by the presence of a magnetic field, and he decided to investigate the effect of a magnetic field on a group of emission lines of the sodium atom known as the sodium D lines. Zeeman decided to investigate the matter by performing a series of measurements at the University of Leiden in 1896.

Zeeman first decided to investigate the effect of the magnetic field on the two emission lines of the sodium atom at wavelengths  $5890.0 \text{ \AA}$  and  $5895.9 \text{ \AA}$ . These two lines are known as the sodium D lines, and they are responsible for the yellow light of a sodium flame. Zeeman observed a distinct widening of the sodium D lines when the magnetic field was turned on, but he could not dissolve the lines. In a subsequent experiment on one of the emission lines of the cadmium atom, he did observe a splitting of the spectral line in the presence of a magnetic field. If the magnetic field is parallel to the direction of observation, the line splits into two components, and if the direction of observation is perpendicular to the magnetic field, there are three components.

The story goes that Zeeman reported his experimental results to *Hendrik Antoon Lorentz* (1853–1928), the professor of theoretical physics at the University of Leiden, and that Lorentz went home and derived the theoretical interpretation of the magnetic splitting of the spectral lines that same evening. Lorentz's theory was based on the classical electron theory. He considered the motion of an electron in a circular orbit, and he calculated the effect of magnetic fields in various directions on the electronic motion. The theory predicted the polarization of the various components, and Zeeman easily verified that the theoretical polarization predictions were consistent with the experimental observations in the cadmium experiment. Lorentz and Zeeman were joint recipients of the second Nobel Prize in physics in 1902. Meanwhile Zeeman was appointed professor of physics at the University of Amsterdam in 1900, and he remained there for the rest of his career.

Unfortunately, in 1897 more precise experiments were performed on the magnetic splittings of the two sodium D lines; it was found that one line splits into four components and the other line splits into six components. This became known as the anomalous Zeeman effect, and it defied all theoretical interpretations. Those splittings that were consistent with Lorentz's theoretical predictions were called the normal Zeeman effect. In spite of the various inconsistencies, the normal Zeeman effect in combination with Lorentz's theoretical interpretation constituted a clear proof of the existence of the electron.

The Stern-Gerlach experiment was successfully performed in February 1922, and it was important because it provided experimental evidence of the quantization of the angular momentum that had been proposed by Ehrenfest and by Sommerfeld

in 1913 (see Section 1.III). It was not immediately understood that it also showed the existence of the electron spin.

As a young man, *Otto Stern* (1888–1969) acquired a broad and thorough knowledge of both theoretical and experimental physics. He was able to do this because he was fortunate enough to have wealthy parents, so that he did not have to worry about earning a living. At first, Stern was primarily interested in theoretical physics, and he worked with Einstein first in Prague and later in Zürich. He received an academic appointment in Frankfurt, but as happened to many of his contemporaries, his career was interrupted by the First World War. When he returned to Frankfurt at the end of the war, he became more interested in experimental physics and made some measurements on atomic beams of silver atoms.

Stern was aware of the quantization rule of Sommerfeld and Ehrenfest, and it occurred to him that its validity could be verified by measuring the effect of a magnetic field on a beam of silver atoms. If the angular momentum was quantized, then it should be possible to split the beam by applying a magnetic field, whereas the absence of quantization would lead to a broadening of the beam by a magnetic field.

The problem was that it was very difficult to perform accurate experiments on atomic silver beams with the available equipment. Fortunately, the physics department at the University of Frankfurt had just succeeded in attracting one of the most able and energetic German experimental physicists to their staff, namely, *Walter Gerlach* (1889–1979).

Stern had already published a detailed outline of the proposed experiment, but it was left to Gerlach to actually implement the idea. This turned out to be a Herculean task that required all of Gerlach's expertise. Einstein helped raise the funds that were necessary to pay for building and upgrading the equipment. Gerlach spent many nights in the laboratory supervising the actual measurements. All these efforts paid off; one morning in February 1922, Gerlach observed unambiguous proof of the expected splitting of the beam of silver atoms. The experimental results were published jointly by Stern and Gerlach in March 1922.

By that time, Stern had left Frankfurt to take up a position as professor of physics at Rostock University. Shortly thereafter he moved to Hamburg University as professor of physical chemistry and director of the Institute for Physical Chemistry. During the next decade, Stern made important contributions to physics in the area of molecular beams. Stern was Jewish, and he resigned his position in 1933 before he could be dismissed by the Nazi regime. He moved to the United States and accepted a position at the Carnegie Institute of Technology in Pittsburg, where he continued his scientific research until he retired in 1946. He was the recipient of the 1943 Nobel Prize in physics, which he received in 1945.

Gerlach remained in Germany. In 1929 he was appointed to the chair of experimental physics at the University of Munich, one of the most prestigious positions in Germany. He was very highly regarded as one of the leading experimental physicists in Germany, and he was elected rector of the University of Munich between 1948 and 1951.

The fourth experimental phenomenon that was relevant to the further development of quantum mechanics was the doublet structure of the emission spectrum

of alkali atoms. We mentioned that Zeeman tried to measure the magnetic splitting of the sodium D lines. The latter consist of a pair of spectral lines in the yellow part of the spectrum with wavelengths separated by 6 Å. It was believed that these splittings were associated with interactions between the orbit of the outer electron and the angular momentum of the core, but it became increasingly difficult to defend this assumption.

The various experimental discoveries outlined above presented a challenge to the theoretical physicists because they could not be explained by making use of classical physics or even the existing quantum theory. They could be explained only after the introduction of the electron spin and of the exclusion principle. We discuss these developments in the following two sections.

### III. PAULI'S EXCLUSION PRINCIPLE

The experimental work on X rays led to a better understanding of atomic structure. We mentioned that the X rays emitted by an atom could be assigned to different groups, namely, K lines, L lines, M lines, and so on. It was found that the emission of a K line was associated with the excitation of the most highly energetic electrons closest to the atomic nucleus that constituted the K shell. The X ray L lines had longer wavelengths and lower energies than the K lines, and they were associated with excitations of electrons in the L shell that were farther away from the nucleus than the K shell electrons.

In 1923 and 1924, Bohr and the Dutch physicist *Dirk Coster* (1889–1950) proposed that the various electron shells could be identified by Bohr's quantum numbers:  $n = 1$  for the K shell,  $n = 2$  for the L shell,  $n = 3$  for the M shell, and so on. This idea was further expanded by *Edmund Clifton Stoner* (1880–1968), who suggested that within a shell the electrons could be further classified according to a second quantum number  $l$ , which can assume the values  $l = 0, 1, 2, \dots, n - 1$ .

Stoner also proposed that the electrons in an atom can all be characterized by a set of the three hydrogen atom quantum numbers  $n$ ,  $l$ , and  $m$ . Here  $n$  represents the energy,  $l$  the magnitude, and  $m$  the orientation of the angular momentum of each electron in the atom (see Section 8.III). It also follows that the various electrons in an atom are generally in different quantum states.

Meanwhile, Pauli had been very interested in the various theoretical problems associated with atomic structure. He was aware of the experimental and theoretical developments, and in 1925 he formulated what became known as the exclusion principle. He assumed that in addition to the three quantum numbers  $n$ ,  $l$ , and  $m$ , each stationary state is characterized by a fourth quantum number that can have only two different values. The exclusion principle states that each atomic stationary state described by the four quantum numbers can accommodate only one electron.

Wolfgang Ernst Pauli was born in April 1900 in Vienna. His father was a distinguished scientist. He was a professor of chemistry at the University of Vienna, director of the Institute for Medical Colloid Chemistry, and was considered one



of the founding fathers of the latter field. Wolfgang's godfather was the equally distinguished physicist Ernst Mach, who became his mentor and advisor.

Pauli was a child prodigy. He began to study physics with Arnold Sommerfeld in Munich, and he published his first scientific paper on relativity theory at age 18. He was not only very intelligent but also highly knowledgeable. In 1921, at age 21, he published a 237-page review article on relativity theory that even today is considered one of the best and most complete texts on the subject.

In 1923 he was appointed a faculty member at the University of Hamburg. It was here that he concluded from a study of the anomalous Zeeman effect that an electron in a stationary state could be in two different sub-states that should be described by a fourth quantum number. Earlier, Stoner had proposed that each atomic shell (K, L, M, etc.) characterized by the quantum number  $n$  could accommodate no more than  $2n^2$  electrons. By combining Stoner's idea with the hypothesis of two-valuedness of the electron, Pauli concluded that in an atom, only one electron could be assigned to the quantum state described by the four quantum numbers  $n$ ,  $l$ ,  $m$ , and  $k$ . The quantum number  $k$  could have only two different values; it represented Pauli's two-valuedness. Pauli was unable to offer any theoretical justification for his exclusion principle. However, it constitutes the basis for all theoretical interpretations of atomic and molecular structure, and it also offers an explanation for the periodic system of the elements.

In 1928 Pauli moved from Hamburg to Zürich, where he became a professor at the famous Eidgenössische Technische Hochschule. It was here that Pauli made his second important contribution to physics: his proposal of the existence of another elementary particle. Pauli remained in Zürich for the remainder of his career except for the period 1940–1945, which he spent at the Institute for Advanced Studies in Princeton.

Pauli's personality exhibited the idiosyncrasies of a genius. He very much enjoyed discussing scientific advances with fellow physicists. He attended numerous physics meetings. He also exchanged correspondence and had private meetings with the majority of the prominent physicists of his time. He was very critical, and he did not always express his criticisms with a great deal of tact. On the other hand, he had no ulterior motives in his criticisms, and he was equally critical about his own accomplishments. On the whole, he was well liked, and he had a major impact on the advance of physics.

During Pauli's travels, experimental physicists noted that experiments often ran into serious problems whenever Pauli came into their vicinity; this became known as the Pauli effect. There was no logical explanation for this effect, but whenever Otto Stern in Hamburg had to meet with his colleague, he always made sure that the meetings were held as far away from his laboratory as possible.

#### IV. THE DISCOVERY OF THE ELECTRON SPIN

The electron spin was discovered in 1925 by *Samuel Abraham Goudsmit* (1902–1978) and *George Eugene Uhlenbeck* (1900–1988), who were both graduate

students at the University of Leiden at that time. Even though they made the discovery entirely on their own, they were greatly helped by the encouragement of their professor, *Paul Ehrenfest* (1880–1933).

When Lorentz retired in 1912 from the chair of theoretical physics at the University of Leiden, he personally selected Ehrenfest as his successor on the recommendation of Sommerfeld. It was in many respects an unusual choice. Ehrenfest had studied in Vienna and Göttingen, but he had never visited the Netherlands. Also, he had never held an academic appointment. He had received his doctorate in Vienna with the famous physicist *Ludwig Eduard Boltzmann* (1844–1906), and since then he had been active in research in theoretical physics. He was well regarded by his peers, but in spite of many efforts he had been unable to secure an academic appointment anywhere. When he was offered the Leiden chair, he resided in St. Petersburg in Russia. He spoke German and Russian but he was not familiar with the Dutch language.

In spite of these drawbacks, Ehrenfest's appointment in Leiden may be considered a success. He made Leiden into a thriving center of theoretical physics by attracting many visitors and students. He was extremely generous and helpful to some of his favorite students. He introduced Goudsmit and Uhlenbeck to each other, and he persuaded them to work on atomic spectra. After they graduated, he found both of them academic positions at the University of Michigan. He was a beloved teacher to his favorite students. On the other hand, he was considerably less friendly to students and assistants whom he considered less talented, and he was not necessarily well liked by those students.

Goudsmit and Uhlenbeck began their joint work on atomic spectra in 1925. Uhlenbeck was Ehrenfest's assistant, while Goudsmit was the assistant of Zeeman in Amsterdam. They were interested in finding a theoretical explanation for the anomalous Zeeman effect and for the doublet structure of the alkali metal spectra, and they were helped in their efforts by some recent developments in those areas.

The German physicist *Alfred Landé* (1888–1976) had proposed in 1921 that most of the aspects of the anomalous Zeeman effect could be explained by assuming that the quantum numbers associated with the angular momentum could assume half-integer values. Landé's proposal is, of course, inconsistent with the original quantization rule

$$M = n\hbar \quad n = 0, 1, 2, \text{ etc.} \quad (10-2)$$

that had been introduced by Ehrenfest and by Sommerfeld in 1913. However, it should be noted that Ehrenfest and Sommerfeld had not offered any fundamental justification for their quantization rule either, other than the fact that it agreed with experimental findings. It could therefore be argued that Landé's suggestion was no less valid than the previous quantization rule.

The doublet structure of the alkali spectra may be explained by assuming that it is due to the interaction between two angular momenta, but it was far from clear what those angular momenta might be. It was first assumed that both the electron

core and the outer electron possessed an angular momentum, but Pauli showed that the core must have zero angular momentum.

Goudsmit and Uhlenbeck were familiar with these developments, and it suddenly occurred to them that the missing angular momentum might actually be attributed to rotational motion of the electron around its axis. The magnitude of the angular momentum of this rotation was assumed to be  $(\hbar/2)$ , so that the two possible projections along a given direction are  $(+\hbar/2)$  and  $(-\hbar/2)$ .

The idea of the spinning electron appears to be quite logical in retrospect. After all, the Earth performs a daily rotation around its axis in addition to a yearly orbit around the sun. The assumption of a similar motion pattern for the electron may therefore not seem unreasonable. Also, it was generally believed that any quantum number is usually related to a degree of freedom, and if Pauli's fourth quantum number had to be associated with some type of motion, then the electronic rotation (or spin) seemed to be the only possibility.

Initially the electron spin hypothesis suffered from a serious discrepancy. The gyromagnetic ratio—that is, the ratio between the spin magnetic moment and its angular momentum—had to be assumed to be twice as large as the corresponding orbital value in order to agree with experimental findings. This difference between the two different gyromagnetic ratios was inconsistent with classical electromagnetic theory, and as a result, Pauli rejected the whole hypothesis of the spinning electron. However, we mentioned in Section 1.VI that the difference between the two gyromagnetic ratios may be explained by taking relativistic effects into account.

Just a few years later, in 1928, Dirac formulated the relativistic wave equation for quantum mechanics. The hypothesis of the electron spin with magnitude  $(\hbar/2)$  is an integral component of the Dirac equation. Dirac's relativistic description was considered the final step in the development of quantum mechanics.

## V. THE MATHEMATICAL DESCRIPTION OF THE ELECTRON SPIN

The most convenient quantum mechanical description of the electron spin is obtained by drawing an analogy with the quantum theory of the angular momentum presented in Section 7.V. In the latter case, we may choose a set of functions  $\psi(l, m)$  that are eigenfunctions of both the operator  $(M^2)_{\text{op}}$  and the operator  $(M_z)_{\text{op}}$ . The eigenvalues of both operators are then defined as follows:

$$\begin{aligned} (M_z)_{\text{op}} \psi(l; m) &= l(l+1)\hbar^2 \psi(l; m) \\ (M_z)_{\text{op}} \psi(l; m) &= m \hbar \psi(l; m) \\ l &= 0, 1, 2, 3, \dots \\ m &= 0, \pm 1, \pm 2, \dots, \pm l \end{aligned} \tag{10-3}$$

We may adopt the same mathematical formalism to represent the electron spin by assuming that the spin quantum number is equal to  $(1/2)$  rather than an integer. The values of the two quantum numbers  $l$  and  $m$  then become

$$\begin{aligned} l &= 1/2 \\ m &= \pm 1/2 \end{aligned} \tag{10-4}$$

The corresponding eigenfunctions are usually denoted by  $\alpha$  and  $\beta$ :

$$\begin{aligned} \psi(1/2, 1/2) &= \alpha \\ \psi(1/2, -1/2) &= \beta \end{aligned} \tag{10-5}$$

The spin functions  $\alpha$  and  $\beta$  satisfy the following orthonormality relations:

$$\begin{aligned} \langle \alpha | \alpha \rangle &= \langle \beta | \beta \rangle = 1 \\ \langle \alpha | \beta \rangle &= \langle \beta | \alpha \rangle = 0 \end{aligned} \tag{10-6}$$

The spin angular momentum operators are denoted by the symbols  $S^2$  and  $S_z$  instead of the symbols  $M^2$  and  $M_z$ , which are reserved for the orbital angular momentum. By analogy with Eq. (10-3) we then have

$$\begin{aligned} (S^2)_{\text{op}}\alpha &= (3\hbar^2/4)\alpha & (S^2)_{\text{op}}\beta &= (3\hbar^2/4)\beta \\ (S_z)_{\text{op}}\alpha &= (\hbar/2)\alpha & (S_z)_{\text{op}}\beta &= (-\hbar/2)\beta \end{aligned} \tag{10-7}$$

We should realize that in the representation (10-3) of the angular momentum, the  $Z$  direction is a preferred direction because the angular momentum is quantized with respect to the  $Z$  axis. This is logical from a mathematical perspective since the polar coordinates are defined relative to the  $Z$  axis and the eigenfunctions of the operator have a particularly simple form in this representation. From a physical point of view, it should make no difference which of the three coordinate axes  $X$ ,  $Y$ , or  $Z$  we select in order to represent the angular momentum. A different choice would still yield the same eigenvalues, but the eigenfunctions would be different.

The effect of a coordinate transformation on the eigenfunctions of the angular momentum may be derived from the matrix representation of the three operators  $(M_x)_{\text{op}}$ ,  $(M_y)_{\text{op}}$ , and  $(M_z)_{\text{op}}$ . We did not derive these matrices in Chapter 7, but in the case of the spin angular momentum they assume a particularly simple form. Rather than present the matrices, we simply report the effect of the spin operators  $(S_x)_{\text{op}}$  and  $(S_y)_{\text{op}}$  on the functions  $\alpha$  and  $\beta$ :

$$\begin{aligned} (S_x)_{\text{op}}\alpha &= (\hbar/2)\beta & (S_x)_{\text{op}}\beta &= (\hbar/2)\alpha \\ (S_y)_{\text{op}}\alpha &= (i\hbar/2)\beta & (S_y)_{\text{op}}\beta &= (-i\hbar/2)\alpha \end{aligned} \tag{10-8}$$

The eigenfunctions of the operators  $(S_x)_{\text{op}}$  and  $(S_y)_{\text{op}}$ , which represent the quantization of the electron spin with respect to the X and Y axes, respectively, are easily derived from Eq. (10-8). We have

$$\begin{aligned}(S_x)_{\text{op}}(\alpha + \beta) &= (\hbar/2)(\alpha + \beta) \\ (S_x)_{\text{op}}(\alpha - \beta) &= (-\hbar/2)(\alpha - \beta)\end{aligned}\tag{10-9}$$

and

$$\begin{aligned}(S_y)_{\text{op}}(\alpha + i\beta) &= (\hbar/2)(\alpha + i\beta) \\ (S_y)_{\text{op}}(\alpha - i\beta) &= (-\hbar/2)(\alpha - i\beta)\end{aligned}\tag{10-10}$$

We now determine the energy eigenvalues of a spinning electron in a homogeneous magnetic field that we denote by  $\mathbf{B}$ . It may be derived from classical electromagnetic theory that for the orbital motion of electrons in an atom, the ratio between the orbital angular momentum  $\mathbf{M}$  and the resulting magnetic moment  $\boldsymbol{\mu}$  is given by

$$\boldsymbol{\mu} = -\frac{e}{2mc} \mathbf{M}\tag{10-11}$$

where  $e$  is the charge of the electron and  $m$  is its mass;  $c$  is the velocity of light.

We mentioned that for a spinning electron the ratio between its spin angular momentum  $\mathbf{S}$  and the corresponding magnetic moment  $\boldsymbol{\mu}_s$  was postulated by Goudsmit and Uhlenbeck to be twice as large as for the corresponding orbital angular momentum and magnetic moment, namely,

$$\boldsymbol{\mu}_s = -\frac{e}{mc} \mathbf{S}\tag{10-12}$$

The assumption of different gyromagnetic ratios for spin and orbital motion was necessary in order to interpret the experimental information. The assumption could not be explained on the basis of classical electromagnetic theory, and this caused Pauli initially to reject the whole idea of the spinning electron. However, both Bohr and Einstein believed that the electron spin hypothesis might have merit because of its excellent agreement with experimental findings and that the difference in gyromagnetic ratios might be due to relativistic effects. Bohr assigned the problem to one of his postdoctoral assistants, Lewellyn Thomas, who proved that the difference in gyromagnetic ratio could indeed be understood on the basis of relativity theory.

The Hamiltonian operator  $H$  of a spinning electron in a magnetic field may now be represented as

$$H = -(\boldsymbol{\mu}_s \cdot \mathbf{B}) = \frac{e}{mc} (B_x S_x + B_y S_y + B_z S_z)\tag{10-13}$$

It follows from Eqs. (10-7), (10-9), and (10-10) that this operator has two eigenvalues

$$E_{1,2} = \pm \frac{e \hbar B}{2mc} \quad (10-14)$$

no matter which direction the magnetic field  $\mathbf{B}$  has. The energy difference between the two eigenvalues is

$$\Delta E = \frac{e \hbar B}{mc} \quad (10-15)$$

This equation constitutes the basis for electric spin resonance (ESR) measurements.

Let us now discuss how the wave function of a bound particle must be represented if the spin is included in our considerations. In general, we should expand the wave function in terms of all possible spin states. In the case of one electron there are only two spin states, characterized by the functions  $\alpha$  and  $\beta$ , so that the general wave function  $\psi(x, y, z, s)$  may always be written as

$$\psi(x, y, z; s) = \psi_+(x, y, z)\alpha + \psi_-(x, y, z)\beta \quad (10-16)$$

If we neglect all relativistic effects and all interactions involving the spin, then the two functions  $\psi_+$  and  $\psi_-$  are eigenfunctions of the same nonrelativistic Hamiltonian  $H_0$ . They should therefore be proportional to one another, and the wave function  $\psi$  may be written as

$$\psi(x, y, z; s) = \phi(x, y, z)[a\alpha + b\beta] \quad (10-17)$$

that is, as a superposition of the two possible spin states. The probability of finding the system in the spin state  $\alpha$  is given by  $a \cdot a^*$  and the probability of a spin  $\beta$  is given by  $b \cdot b^*$ .

## VI. THE EXCLUSION PRINCIPLE REVISITED

The initial formulation of the exclusion principle by Pauli discussed in Section 10.III stated that in an atom no more than one electron could be assigned the same set of four quantum numbers  $n$ ,  $l$ ,  $m$ , and  $k$ . The fourth quantum number,  $k$ , was subsequently identified with the spin quantum number ( $\pm 1/2$ ) of the electron. This description of the exclusion principle was limited in its scope. It was restricted to atoms, and it was based on the assumption that the electrons could all be identified by hydrogen-like quantum numbers.

It is not surprising that Pauli's original definition of the exclusion principle was soon replaced by a more general formulation that applies to all many-electron systems. The latter description makes use of the mathematical concept of permutations discussed in Section 2.V. We consider an  $N$ -electron system that is represented by a Hamiltonian operator  $H(1, 2, 3, \dots, N)$ , and we note that in all known atomic and molecular systems this Hamiltonian is symmetric with respect to permutations of the electrons

$$PH(1, 2, \dots, N) = H(1, 2, \dots, N)P \quad (10-18)$$

This is equivalent to stating that  $P$  and  $H$  commute:

$$PH = HP \quad (10-19)$$

We showed in Section 7.II that commuting operators have common eigenfunctions. Any eigenfunction  $\Phi_n$  of the operator  $H$  is therefore also an eigenfunction of the operator  $P$ .

The exclusion principle now requires that any eigenfunction of a many-electron system must be antisymmetric with respect to permutations of the electrons:

$$P\Psi_n(1, 2, \dots, N) = -\Psi_n(1, 2, \dots, N) \quad (10-20)$$

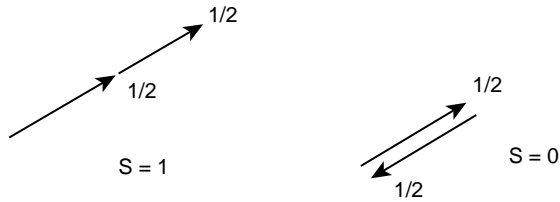
It should be noted that the eigenfunction  $\Psi_n$  must also include the spin coordinates of the electrons.

It is easily seen that the more general definition (10-20) of the exclusion principle also includes Pauli's original formulation. If an atomic configuration were to include two different electrons with the same four quantum numbers, then the corresponding wave function would be symmetric with respect to a permutation of those two electrons. The latter contradicts the requirement of Eq. (10-20), and the configuration is therefore not allowed.

## VII. TWO-ELECTRON SYSTEMS

The wave function of a two-electron system contains a spin-dependent and an orbital-dependent part. We first consider the spin-dependent part since it may be derived by using a simple vector model for the addition of angular momenta.

In a two-electron system each electron has a spin with quantum number  $s = 1/2$ . We may now take the spin angular momentum of one electron as the axis of quantization for the second electron spin (Figure 10-1). The second electron spin has then two possible orientations. It can either point in the same direction as the first electron spin to give a total spin  $S = 1$  or it can point in the opposite direction, which results in a total spin  $S = 0$ . It is then customary to denote the possible



**Figure 10-1** Addition of two spin angular momentum vectors of magnitude 1/2.

projections of the total spin on the Z axis by a second quantum number  $m_S$ . In this way we obtain the following four spin functions:

$$\begin{array}{lll}
 \psi(1, 1) & S = 1 & m_S = 1 \\
 \psi(1, 0) & S = 1 & m_S = 0 \\
 \psi(1, -1) & S = 1 & m_S = -1 \\
 \psi(0, 0) & S = 0 & m_S = 0
 \end{array} \tag{10-21}$$

The spin state  $S = 1$  is threefold degenerate, and it is generally known as a triplet state, whereas the spin state  $S = 0$  is called a singlet state because it is nondegenerate.

It is now relatively easy to derive the detailed form of the four spin functions of Eq. (10-21). The total spin operator  $S$  is defined as

$$\begin{aligned}
 \mathbf{S} &= \mathbf{S}_1 + \mathbf{S}_2 \\
 S_z &= S_{1,z} + S_{2,z} \dots \text{ etc.}
 \end{aligned} \tag{10-22}$$

The four possible spin functions are all possible products of the two spin functions  $\alpha$  and  $\beta$ . Operating  $(S_z)_{\text{op}}$  on the four function gives the following results:

$$\begin{aligned}
 (S_z)_{\text{op}}\alpha(1)\alpha(2) &= (S_{1,z} + S_{2,z})_{\text{op}}\alpha(1)\alpha(2) = \hbar\alpha(1)\alpha(2) \\
 (S_z)_{\text{op}}\beta(1)\beta(2) &= (S_{1,z} + S_{2,z})_{\text{op}}\beta(1)\beta(2) = -\hbar\beta(1)\beta(2) \\
 (S_z)_{\text{op}}\alpha(1)\beta(2) &= (S_{1,z} + S_{2,z})_{\text{op}}\alpha(1)\beta(2) = 0 \\
 (S_z)_{\text{op}}\beta(1)\alpha(2) &= (S_{1,z} + S_{2,z})_{\text{op}}\beta(1)\alpha(2) = 0
 \end{aligned} \tag{10-23}$$

It follows immediately that

$$\begin{aligned}
 \psi(1, 1) &= \alpha(1)\alpha(2) \\
 \psi(1, -1) &= \beta(1)\beta(2)
 \end{aligned} \tag{10-24}$$



and it is easily derived that the other two spin functions are given by

$$\begin{aligned}\psi(1, 0) &= (1/\sqrt{2})[\alpha(1)\beta(2) + \beta(1)\alpha(2)] \\ \psi(0, 0) &= (1/\sqrt{2})[\alpha(1)\beta(2) - \beta(1)\alpha(2)]\end{aligned}\tag{10-25}$$

We see that the three triplet spin functions are all symmetric with respect to permutations of the two electrons, while the singlet spin function is antisymmetric.

The total wave function  $\Psi(1, 2)$  of a two-electron system is now represented as a product of an orbital function  $\Phi(\mathbf{r}_1, \mathbf{r}_2)$  and a spin function  $\psi(s, m_s)$

$$\Psi(1, 2) = \Phi(\mathbf{r}_1, \mathbf{r}_2)\psi(s, m_s)\tag{10-26}$$

According to the exclusion principle, this function should be antisymmetric with respect to permutations:

$$\Psi(2, 1) = -\Psi(1, 2)\tag{10-27}$$

It is easily seen that for a singlet state the orbital function is symmetric with respect to permutations of the orbital coordinates  $\mathbf{r}_1$  and  $\mathbf{r}_2$ , and for a triplet state the function is antisymmetric with respect to such permutations. The reverse of this statement is also of interest. If the orbital function is symmetric it must correspond to a singlet spin state, and if it is antisymmetric it must be a triplet state.

## VIII. THE HELIUM ATOM

It is now generally accepted that the electrons in an atom may be identified by a set of hydrogen-type quantum numbers. An atomic eigenstate may then be characterized by its electronic configuration, which describes the assignment of the electrons. The lowest eigenstate of the helium atom has both electrons in the  $(1s)$  state and it is called  $(1s)^2$ . Excited states are obtained by exciting one of the two electrons to a higher eigenstate; examples are  $(1s)(2s)$ ,  $(1s)(2p)$ ,  $(1s)(3s)$ , and so on. Eigenstates in which both electrons are excited are of little interest.

In this approach, it is further assumed that the atomic wave function may be written as a product of one-electron wave functions known as orbitals. This approximation is called the Hartree-Fock approximation. In constructing the atomic wave function, both the exclusion principle and the existence of the electron spins must be taken into account. This means that the atomic wave function assumes a more complex form than just a simple product.

We denote the lowest orbital of the helium atom by  $\phi_1$ . Then the corresponding orbital wave function is

$$\Phi(\mathbf{r}_1, \mathbf{r}_2) = \phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2)\tag{10-28}$$

This function is symmetric with respect to permutations, so it must be multiplied by an antisymmetric spin function in order to satisfy the exclusion principle:

$$\Psi_1(1, 2) = (1\sqrt{2})\phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2)[\alpha(1)\beta(2) - \beta(1)\alpha(2)] \quad (10-29)$$

An alternative formulation is

$$\Psi_1(1, 2) = (1\sqrt{2}) \sum P\delta_p[\phi_1(\mathbf{r}_1)\alpha(1)\phi_1(\mathbf{r}_2)\beta(2)] \quad (10-30)$$

It follows that a configuration with two electrons in identical orbitals must necessarily be a singlet state.

We now consider a different configuration where one electron is in an orbital  $\phi_1$  and the second electron is in a different orbital  $\phi_2$ . From the product of the two orbitals we can construct two different orbital functions that are symmetric or antisymmetric with respect to permutations:

$$\begin{aligned} \Phi_s(\mathbf{r}_1, \mathbf{r}_2) &= (1\sqrt{2})[\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) + \phi_2(\mathbf{r}_1)\phi_1(\mathbf{r}_2)] \\ \Phi_a(\mathbf{r}_1, \mathbf{r}_2) &= (1\sqrt{2})[\phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2) - \phi_2(\mathbf{r}_1)\phi_1(\mathbf{r}_2)] \end{aligned} \quad (10-31)$$

The symmetric function should obviously be combined with an antisymmetric singlet spin function

$${}^1\Psi(1, 2) = (1\sqrt{2})\Phi_s(\mathbf{r}_1, \mathbf{r}_2)[\alpha(1)\beta(2) - \beta(1)\alpha(2)] \quad (10-32)$$

and the antisymmetric function with a symmetric triplet function, for example,

$${}^3\Psi(1, 2) = \Phi_a(\mathbf{r}_1, \mathbf{r}_2)\alpha(1)\alpha(2) \quad (10-33)$$

It follows that the spin multiplicity determines the symmetry of the orbital wave function. Consequently, singlet and triplet states belonging to the same electronic configuration may have very different energies.

The helium atom Hamiltonian is given by

$$H = -\frac{1}{2}\Delta_1 - \frac{1}{2}\Delta_2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}} \quad (10-34)$$

if we use the atomic units of length and energy  $a_0$  and  $\epsilon_0$  defined in Eqs. (8-22) and (8-23). It is convenient to write this Hamiltonian as

$$H = G(1) + G(2) + \Omega(1, 2) \quad (10-35)$$

where

$$G(i) = -\frac{1}{2}\Delta_i + \frac{2}{r_i} \quad (10-36)$$

$$\Omega(i,j) = \frac{1}{r_{i,j}}$$

The energy expectation value of the helium atom ground state may then be derived from Eqs. (10-29) and (10-35). It is given by

$$\langle \Psi | H | \Psi \rangle = 2\langle \phi_1 | G | \phi_1 \rangle + \langle \phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2) | \Omega(1,2) | \phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2) \rangle \quad (10-37)$$

since the integral of the singlet spin function is

$$\langle \alpha(1)\beta(2) - \beta(1)\alpha(2) | \alpha(1)\beta(2) - \beta(1)\alpha(2) \rangle = 2 \quad (10-38)$$

according to Eq. (10-6). We may abbreviate Eq. (10-37) as

$$\langle \Psi | H | \Psi \rangle = 2G_1 + J_{1,1} \quad (10-39)$$

by defining the one-electron integrals

$$G_i = \langle \phi_i | G | \phi_i \rangle \quad (10-40)$$

and the so-called Coulomb integral

$$J_{k,l} = \langle \phi_k(\mathbf{r}_1)\phi_l(\mathbf{r}_2) | \Omega(1,2) | \phi_k(\mathbf{r}_1)\phi_l(\mathbf{r}_2) \rangle \quad (10-41)$$

The energies of the excited singlet and triplet configurations may be derived in a similar fashion from the wave functions (10-32) and (10-33). We find then that

$$\begin{aligned} \langle {}^1\Psi(1,2) | H | {}^1\Psi(1,2) \rangle &= \langle \Phi_s(\mathbf{r}_1, \mathbf{r}_2) | H | \Phi_s(\mathbf{r}_1, \mathbf{r}_2) \rangle \\ &= G_1 + G_2 + J_{1,2} + K_{1,2} \end{aligned} \quad (10-42)$$

for the singlet state and

$$\begin{aligned} \langle {}^3\Psi(1,2) | H | {}^3\Psi(1,2) \rangle &= \langle \Phi_a(\mathbf{r}_1, \mathbf{r}_2) | H | \Phi_s(\mathbf{r}_1, \mathbf{r}_2) \rangle \\ &= G_1 + G_2 + J_{1,2} - K_{1,2} \end{aligned} \quad (10-43)$$

for the corresponding triplet state. Here the integral  $K_{1,2}$  is known as an exchange integral; it is defined as

$$K_{k,l} = \langle \phi_k(\mathbf{r}_1)\phi_l(\mathbf{r}_2) | \Omega(1,2) | \phi_l(\mathbf{r}_1)\phi_k(\mathbf{r}_2) \rangle \quad (10-44)$$

It follows that for a configuration  $(\phi_1)(\phi_2)$  the energies of the corresponding singlet and triplet states differ by an amount that is equal to twice the exchange integral. Since the latter is positive, the triplet state will have the lower energy. It may also be seen that the value of the exchange integral is easily derived from the energy difference between the triplet and singlet states.

## IX. THE HELIUM ATOM ORBITALS

It is possible to obtain approximate expressions for the helium atom orbitals by drawing an analogy with the hydrogen atom. We first consider the orbital  $\phi_1$  that corresponds to the ground state configuration  $(1s)^2$  of the helium atom. Here, each of the two electrons moves in the potential field of the nucleus and of the other electron. If we assume that the charge cloud of the second electron is spherically symmetric, we may use classical electrostatic theory to get a rough idea of what this potential field looks like. We define  $\delta(r)$  as the part of the probability density of either one of the two electrons that is contained in a sphere of radius  $r$  around the nucleus. It is defined as

$$\delta(r) = 4\pi \int_0^r [\phi_1(r)]^2 r^2 dr \quad (10-45)$$

Here it is assumed that  $\phi_1$  depends on the variable  $r$  only and that it is normalized to unity. The potential field of the other electron is then given by

$$V(r) = -\frac{2 - \delta(r)}{r} = -\frac{Z(r)}{r} \quad (10-46)$$

We know that

$$\begin{aligned} Z(r) &= 2 & \text{if } r &= 0 \\ Z(r) &= 1 & \text{if } r &\rightarrow \infty \end{aligned} \quad (10-47)$$

and we assume that we may replace it by an effective average nuclear charge  $Z$ . The corresponding orbital is then a hydrogen ( $1s$ ) orbital corresponding to a charge  $Z$ , namely,

$$\phi_1(r) = s_0(r) = (Z^3/\pi)^{1/2} \exp(-Zr) \quad (10-48)$$

We know that

$$\left(-\frac{1}{2}\Delta - \frac{Z}{r}\right)s_0(r) = -\frac{Z^2}{2}s_0(r) \quad (10-49)$$

because  $s_0$  is an eigenfunction of a Coulomb field with charge  $Z$ . It is then easily derived that

$$\langle s_0 | G | s_0 \rangle = -\frac{Z^2}{2} + Z(Z - 2) = \frac{Z^2}{2} - 2Z \quad (10-50)$$

It may also be shown that the Coulomb integral  $J$  is

$$\langle s_0(r_1)s_0(r_2) | \Omega(1, 2) | s_0(r_1)s_0(r_2) \rangle = (5Z/8) \quad (10-51)$$

Substitution into Eq. (10-39) leads to the following expression for the expectation value of the energy:

$$\langle E \rangle = Z^2 - 4Z + \frac{5Z}{8} = Z^2 - \frac{27Z}{8} \quad (10-52)$$

According to the variational principle, the best possible wave function is obtained by minimizing Eq (10-52) with respect to  $Z$ . The results are

$$\begin{aligned} Z_{\min} &= 1.6875 \\ E_{\min} &= -2.847656 \text{ hartree} \end{aligned} \quad (10-53)$$

The experimental ground state energy of the helium atom is  $E = -2.90372$  hartree, so our approximate result has an error of about 2%. We may also conclude that the shielding of the nuclear charge 2 by the other electron amounts to about 30% of the charge of the electron.

Much more complex and sophisticated variational functions for the helium atom have been proposed and minimized, and in this way highly accurate energy values have been calculated. In fact, the ground state energy has been reproduced to within the experimental error. On the other hand, no exact solution of the helium atom Schrödinger equation has been obtained, so far and most physicists believe that this problem will never be solved.

## X. CONCLUDING REMARKS

Many of the general ideas and concepts presented in this chapter are important for the quantum mechanics of many electron systems. We believed that it might be helpful to first discuss some of their applications for a two-electron system such as the helium atom because this enabled us to present the mathematics in a more detailed manner. For instance, we were able to present explicit general expressions for the orbital and spin functions that are consistent with the exclusion principle.

We shall see that in discussing the quantum mechanics of larger atoms and molecules we are often limited to more general mathematical representations. Also,

many-electron systems are subject to additional complexities that are not encountered in two-electron systems. For these reasons, we believe it preferable to consider the helium atom separately before proceeding to a general discussion of larger systems.

## XI. PROBLEMS

- 10-1** Electron spin resonance is performed with electromagnetic radiation of 3.15 cm wavelength. Determine the magnitude of the magnetic field for which the separation of the two energy levels of a free electron corresponds to radiation of the above wave length.
- 10-2** Determine the energy eigenvalues and the corresponding spin eigenfunctions of a free electron in a magnetic field  $\mathbf{B}$  in an arbitrary direction. The spin Hamiltonian  $H_s$  is given by

$$H_s = (e \hbar / mc)(\mathbf{B} \cdot \mathbf{S})$$

- 10-3** Derive the four different spin eigenfunctions of the operators  $S^2$  and  $S_z$  for a two-electron system. The spin operator is defined as

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$$

- 10-4** Derive the energy eigenvalues of a two-electron triplet state in a magnetic field  $\mathbf{B}$  directed along the Z axis.
- 10-5** Derive the energy eigenvalues and the corresponding spin eigenfunctions of a two-electron triplet state in a magnetic field  $\mathbf{B}$  in an arbitrary direction. Show that the results are consistent with the results of problem 10-4.
- 10-6** An approximate result for the energy of the  $(1s)(2s)$  configuration of the helium atom may be derived on the assumption that the  $(1s)$  electron experiences a nuclear charge  $Z = 2$  and that its orbital may be approximated by  $\phi(1s) = \exp(-2r)$  while the  $(2s)$  electron experiences a nuclear charge  $Z = 1$  and its orbital may be approximated as  $\phi(1s) = r \exp(-r/2)$ .
- a) Construct a set of orthonormal orbitals  $\phi'(1s)$  and  $\phi'(2s) = \lambda\phi(1s) - \mu\phi(2s)$  that satisfy the conditions

$$\begin{aligned} \langle \phi'(1s) | \phi'(1s) \rangle &= \langle \phi'(2s) | \phi'(2s) \rangle = 1 \\ \langle \phi'(1s) | \phi'(2s) \rangle &= 0 \end{aligned}$$

- b) Construct antisymmetrized singlet and triplet wave functions of the helium atom from the above orbitals.

- c) Express the helium atom energies of the  $(1s)(2s)$  singlet and triplet configurations in terms of integrals containing the orbitals  $\phi'(1s)$  and  $\phi'(2s)$ .

**10-7** Evaluate the integral

$$I = \iint r_{12}^{-1} \exp(-r_1 - r_2) d\mathbf{r}_1 d\mathbf{r}_2$$

by means of conventional integration methods. The integral can be calculated by first integrating over the coordinates of electron 2 by introducing polar coordinates with the vector  $\mathbf{r}_1$  as reference axis.

---

# 11

---

## ATOMIC STRUCTURE

### I. INTRODUCTION

The exclusion principle was not only important in quantum mechanics but also led to a better understanding of chemical principles by supplying an explanation of the Aufbau principle. The latter may be translated as the building-up principle of the elements, and it helps to explain the periodic system.

We have seen that the hydrogen atom consists of a positively charged nucleus surrounded by a negatively charged electron, both with charge  $e$ . It may therefore be considered the smallest or simplest atom. According to the Aufbau principle, the sequence of atoms belonging to the different elements may now be obtained by adding one electron at a time to each previous atom and by increasing its nuclear charge by an amount  $e$ . An atom is then characterized by its atomic number  $Z$ , which defines its place in the sequence; it has  $Z$  electrons and the nuclear charge  $Ze$ .

Each electron in an atom is identified by its orbital, which is defined by a set of hydrogen-like quantum numbers  $(n, l, m)$  and a fourth quantum number representing the spin. In the atomic ground state the electrons should be assigned to the orbitals with the lowest energies, but this assignment should be consistent with the exclusion principle. Accordingly, the hydrogen atom ( $Z = 1$ ) has a configuration  $(1s)$  and the next atom, helium ( $Z = 2$ ), has a configuration  $(1s)^2$ , but the following atom lithium ( $Z = 3$ ) must have a configuration  $(1s)^2 (2s)$  since we may place only



two electrons in the orbital with the lowest energy, namely, the  $1s$  orbital. It may be seen that it is necessary to obey the exclusion principle in implementing the Aufbau principle.

It is also important to know the relative energies of the various orbitals. In the case of the hydrogen atom, the energy of each eigenstate depends only on the quantum number  $n$ , while the quantum numbers  $l$  and  $m$  determine the magnitude and direction of the angular momentum. However, in the case of the other atoms, the orbital energies depend mainly on the quantum number  $n$ , but they also depend to a lesser extent on the value of the quantum number  $l$ .

It is possible to rank the orbital energies according to their relative magnitudes as follows:

$$\begin{aligned} \epsilon(1s) < \epsilon(2s) < \epsilon(2p) < \epsilon(3s) < \epsilon(3p) < \epsilon(4s) < \epsilon(3d) < \epsilon(4p) \\ < \epsilon(5s) < \epsilon(4d) < \epsilon(5p) < \dots, \text{ etc.} \end{aligned} \quad (11-1)$$

This scheme offers a guideline for the assignment of the electrons in atomic ground states, and it is essential for the implementation of the Aufbau principle. In general, the orbital energies depend more strongly on the quantum number  $n$  than on the quantum number  $l$ , but it may be seen that the dependence on  $l$  increases with increasing values of both quantum numbers. The result is that the energy  $\epsilon(4s)$  is actually lower than  $\epsilon(3d)$ ,  $\epsilon(5s)$  is lower than  $\epsilon(4d)$ , and so on. We might add that Eq. (11-1) is consistent with experimental information with regard to the known atomic ground state electronic configurations.

In order to illustrate the Aufbau principle, we have listed the ground-state electronic configurations of the first 36 elements in Table 11-1. It may be seen that many of these configurations represent degenerate or near-degenerate atomic eigenstates. For instance, the ground-state configuration of the nitrogen atom contains three electrons in the  $(2p)$  orbitals. Since the  $(2p)$  energy level is threefold degenerate, there is more than one way that the three electrons may be distributed over these three states, and it follows that the  $(2p)^3$  configuration corresponds to two different atomic eigenstates with slightly different energies. According to a general rule first formulated by the German physicist *Friedrich Hund* (1896–1997), the atomic eigenstate with the lowest energy is the state with the largest value of the total spin angular momentum  $S$  and, for a given  $S$  value, the state with the largest value of the orbital angular momentum  $L$ .

The validity of Hund's rule may be verified in Table 11-1 since we have listed the symbols that describe the values of the various angular momentum vectors in the atomic ground states. The capital letters  $S$ ,  $P$ ,  $D$ ,  $F$ , and so on describe the value of the total orbital angular momentum  $L$ , with  $S$  denoting  $L = 0$ ,  $P$  denoting  $L = 1$ ,  $D$  denoting  $L = 2$ ,  $F$  denoting  $L = 3$ , and so on. The superscript on the left refers to the spin angular momentum; the superscript is equal to the spin multiplicity  $2S + 1$ , where  $S$  is the magnitude of the total spin angular momentum. The value of the total atomic angular momentum  $J$  is described by the subscript on the right; the value of  $J$  depends on the relative orientations of the vectors  $\mathbf{L}$  and  $\mathbf{S}$ .

**Table 11-1. Ground-State Configurations of Selected Elements**

1	H	$^2S_{1/2}$	(1s)
2	He	$^1S_0$	(1s) <sup>2</sup>
3	Li	$^2S_{1/2}$	(1s) <sup>2</sup> (2s)
4	Be	$^1S_0$	(1s) <sup>2</sup> (2s) <sup>2</sup>
5	B	$^2P_{1/2}$	(1s) <sup>2</sup> (2s) <sup>2</sup> (2p)
6	C	$^3P_0$	(1s) <sup>2</sup> (2s) <sup>2</sup> (2p) <sup>2</sup>
7	N	$^4S_{3/2}$	(1s) <sup>2</sup> (2s) <sup>2</sup> (2p) <sup>3</sup>
8	O	$^3P_2$	(1s) <sup>2</sup> (2s) <sup>2</sup> (2p) <sup>4</sup>
9	F	$^2P_{3/2}$	(1s) <sup>2</sup> (2s) <sup>2</sup> (2p) <sup>5</sup>
10	Ne	$^1S_0$	(1s) <sup>2</sup> (2s) <sup>2</sup> (2p) <sup>6</sup>
11	Na	$^2S_{1/2}$	... (2s) <sup>2</sup> (2p) <sup>6</sup> (3s)
12	Mg	$^1S_0$	... (2s) <sup>2</sup> (2p) <sup>6</sup> (3s) <sup>2</sup>
13	Al	$^2P_{1/2}$	... (2s) <sup>2</sup> (2p) <sup>6</sup> (3s) <sup>2</sup> (3p)
14	Si	$^3P_0$	... (2s) <sup>2</sup> (2p) <sup>6</sup> (3s) <sup>2</sup> (3p) <sup>2</sup>
15	P	$^4S_{3/2}$	... (2s) <sup>2</sup> (2p) <sup>6</sup> (3s) <sup>2</sup> (3p) <sup>3</sup>
16	S	$^3P_2$	... (2s) <sup>2</sup> (2p) <sup>6</sup> (3s) <sup>2</sup> (3p) <sup>4</sup>
17	Cl	$^2P_{3/2}$	... (2s) <sup>2</sup> (2p) <sup>6</sup> (3s) <sup>2</sup> (3p) <sup>5</sup>
18	A	$^1S_0$	... (2s) <sup>2s</sup> (2p) <sup>6</sup> (3s) <sup>2</sup> (3p) <sup>6</sup>
19	K	$^2S_{1/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (4s)
20	Ca	$^1S_0$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (4s) <sup>2</sup>
21	Sc	$^2D_{3/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d)(4s) <sup>2</sup>
22	Ti	$^3F_2$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>2</sup> (4s) <sup>2</sup>
23	V	$^4F_{3/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>3</sup> (4s) <sup>2</sup>
24	Cr	$^7S_3$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>5</sup> (4s)
25	Mn	$^6S_{5/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>5</sup> (4s) <sup>2</sup>
26	Fe	$^5D_4$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>6</sup> (4s) <sup>2</sup>
27	Co	$^4F_{9/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>7</sup> (4s) <sup>2</sup>
28	Ni	$^3F_4$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>8</sup> (4s) <sup>2</sup>
29	Cu	$^2S_{1/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>10</sup> (4s)
30	Zn	$^1S_0$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>10</sup> (4s) <sup>2</sup>
31	Ga	$^2P_{1/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>10</sup> (4s) <sup>2</sup> (4p)
32	Ge	$^3P_0$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>10</sup> (4s) <sup>2</sup> (4p) <sup>2</sup>
33	As	$^4S_{3/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>10</sup> (4s) <sup>2</sup> (4p) <sup>3</sup>
34	Se	$^3P_2$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>10</sup> (4s) <sup>2</sup> (4p) <sup>4</sup>
35	Br	$^2P_{3/2}$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>10</sup> (4s) <sup>2</sup> (4p) <sup>5</sup>
36	Kr	$^1S_0$	... (3s) <sup>2</sup> (3p) <sup>6</sup> (3d) <sup>10</sup> (4s) <sup>2</sup> (4p) <sup>6</sup>

It should be noted that the majority of stable atoms or molecules have closed-shell ground states where each occupied orbital contains a pair of electrons. The closed-shell configurations correspond therefore to singlet spin states. There are a few exceptions; a well-known example is the oxygen molecule, which has a triplet ground state.

## II. ATOMIC AND MOLECULAR WAVE FUNCTION

In Section 10.VIII we presented analytical expressions for the wave functions of two-electron systems based on the assumption that these wave functions may be approximated as antisymmetrized products of one-electron orbitals. We use this same approach to derive expressions for the wave functions of many-electron systems.

We first consider the simplest possible case of a closed-shell ground state where we have an even number  $2N$  of electrons and where we place a pair of electrons in each of the orbitals  $\phi_1, \phi_2, \phi_3, \dots, \phi_N$ . The properly antisymmetrized wave function, including spin, is then given by

$$\Psi_0 = [1/(2N)!]^{1/2} \sum_P P \delta_P [\phi_1(\mathbf{r}_1)\alpha(1)\phi_1(\mathbf{r}_2)\beta(2)\phi_2(\mathbf{r}_3)\alpha(3)\phi_2(\mathbf{r}_4)\beta(4) \dots \phi_N(\mathbf{r}_{2N-1})\alpha(2N-1)\phi_N(\mathbf{r}_{2N})\beta(2N)] \quad (11-2)$$

This expression may also be represented in abbreviated form as

$$\Psi_0 = [1/(2N)!]^{1/2} \sum_P P \delta_P \left[ \prod_{i=1}^N \phi_i(\mathbf{r}_{2i-1})\alpha(2i-1)\phi_i(\mathbf{r}_{2i})\beta(2i) \right] \quad (11-3)$$

*John Clarke Slater* (1900–1976) first noted that the expression (11-2) for the antisymmetrized wave function of a closed-shell state is identical to the definition (2-42) of a determinant, and he concluded that the wave function may therefore also be written as a determinant. These determinants became known as Slater determinants, and their use became fairly popular. However, in our experience, the determinant representation does not offer any advantages over the conventional definition (11-2).

It is easily seen that the total spin corresponding to a closed-shell atomic or molecular configuration should be zero since the configuration consists of electron pairs with opposite spins. We will now consider singly excited configurations of the type

$$(\phi_1)^2(\phi_2)^2 \dots (\phi_{j-1})^2(\phi_j)(\phi_{j+1})^2 \dots (\phi_N)^2(\phi_n) \quad (11-4)$$

where one of the electrons is excited from a lower filled orbital  $\phi_j$  to a higher, previously unfilled orbital  $\phi_n$ . This configuration may correspond either to a singlet or to a triplet spin state. We denote the corresponding wave functions by  $\Psi(j \rightarrow n)$ .

We first derive the mathematical expressions from Eq. (11-2), and we assume then that the electron is excited from the orbital  $\phi_N$  to a higher orbital  $\phi_n$ . We find that the wave function of the corresponding excited singlet configuration

${}^1\Psi(N \rightarrow n)$  is given by

$$\begin{aligned}
 {}^1\Psi(N \rightarrow n) = & [1/\{2 \cdot (2N)!\}]^{1/2} \sum_P P\delta_P [\phi_1(\mathbf{r}_1)\alpha(1)(\phi_1)(\mathbf{r}_2)\beta(2)\phi_2(\mathbf{r}_3)\alpha(3) \\
 & \phi_2(\mathbf{r}_4)\beta(4) \cdots \phi_{N-1}(\mathbf{r}_{2N-3})\alpha(2N-3)\phi_{N-1}(\mathbf{r}_{2N-2})\beta(2N-2) \\
 & \{\phi_N(\mathbf{r}_{2N-1})\phi_n(\mathbf{r}_{2N}) + \phi_n(\mathbf{r}_{2N-1})\phi_N(\mathbf{r}_{2N})\}\alpha(2N-1)\beta(2N)] \quad (11-5)
 \end{aligned}$$

whereas one of the corresponding triplet function is given by

$$\begin{aligned}
 {}^3\Psi(N \rightarrow n) = & [1/(2N)!]^{1/2} \sum_P P\delta_P [\phi_1(\mathbf{r}_1)\alpha(1)(\phi_1)(\mathbf{r}_2)\beta(2)\phi_2(\mathbf{r}_3)\alpha(3) \\
 & \phi_2(\mathbf{r}_4)\beta(4) \cdots \phi_{N-1}(\mathbf{r}_{2N-3})\alpha(2N-3)\phi_{N-1}(\mathbf{r}_{2N-2})\beta(2N-2) \\
 & \phi_N(\mathbf{r}_{2N-1})\phi_n(\mathbf{r}_{2N})\alpha(2N-1)\alpha(2N)] \quad (11-6)
 \end{aligned}$$

These expressions are easily generalized to  ${}^1\Psi(j \rightarrow n)$  and  ${}^3\Psi(j \rightarrow n)$ , but we do not present the detailed forms of the corresponding functions.

If an atomic configuration is degenerate in its first approximation, then its wave function must be represented as a linear combination of more than one function of the type of (11-3) or of more than one Slater determinant. The mathematical representation of these situations becomes more complex, and it falls outside the scope of this book.

We will make use of Eq. (11-2) for the wave function of a closed-shell atomic or molecular configuration to derive the Hartree-Fock equations in the next section.

### III. THE HARTREE-FOCK METHOD

The Hartree-Fock approach is based on the assumption that a molecular or atomic wave function may be approximated as an antisymmetrized product of one-electron orbitals in the case of a closed-shell configuration or as a number of antisymmetrized products in other cases. The goal of the Hartree-Fock method or the Self-Consistent Field (SCF) method is the subsequent derivation of the best possible one-electron orbitals by making use of the variational principle.

Rather than vary each orbital at a time, the Hartree-Fock equations have been designed in such a way that all orbitals may be obtained at the same time as the eigenfunctions of an effective one-electron operator, the Hartree-Fock operator. The latter operator contains the average Coulomb repulsion between the electrons. The Hartree-Fock equations for most atoms may be solved exactly in numerical form, but in the case of molecules it is usually necessary to introduce additional approximations.

In order to derive the Hartree-Fock equations, we must first derive the expectation values of the Hamiltonian expressed in terms of one-electron orbitals from the antisymmetrized wave function (11-2). The Hartree-Fock equations are then

obtained by varying one of the one-electron orbitals followed by a number of mathematical transformations. We limit ourselves to a discussion of closed-shell ground state configurations. Other configurations lead to similar sets of equations, but their derivations are more complex.

In order to derive the Hartree-Fock equations, we write the Hamiltonian as a sum of one-electron and two-electron terms:

$$H = \sum_j G(j) + \sum_{j>k} \Omega(j, k) \quad (11-7)$$

Here the one-electron terms  $G(j)$  represent the sum of the kinetic and potential energy of each electron and the two-electron terms are the Coulomb repulsion energies between the electrons. The expectation value  $E$  of the energy is then given by

$$E = \left\langle \Psi_0 \left| \sum_j G(j) + \sum_{j>k} \Omega(j, k) \right| \Psi_0 \right\rangle \quad (11-8)$$

where  $\Psi_0$  is defined in Eq. (11-2). We assume that the one-electron orbitals are orthonormal,

$$\begin{aligned} \langle \phi_j | \phi_k \rangle &= 0 & \text{if } j \neq k \\ \langle \phi_j | \phi_k \rangle &= 1 & \text{if } j = k \end{aligned} \quad (11-9)$$

It is then easily verified that

$$\langle \Psi_0 | \Psi_0 \rangle = 1 \quad (11-10)$$

There is no need to consider all possible permutations in both functions  $\Psi_0$  that occur in Eq. (11-8), the permutations in one of them suffices. If in addition we separate the Hamiltonian into two parts then Eq. (11-8) may be reduced to

$$\begin{aligned} E = & \langle \phi_1(1)\alpha(1)\phi_1(2)\beta(2)\phi_2(3)\alpha(3)\phi_2(4)\beta(4) \cdots \phi_N(2N-1)\alpha(2N-1)\phi_N(2N)\beta(2N) \\ & \left| \sum_j G(j) + \sum_{k>j} \Omega(j, k) \right| \sum_P P \delta_P [\phi_1(1)\alpha(1)\phi_1(2)\beta(2)\phi_2(3)\alpha(3)\phi_2(4)\beta(4) \cdots \\ & \phi_N(2N-1)\alpha(2N-1)\phi_N(2N)\beta(2N)] \rangle \end{aligned} \quad (11-11)$$

We first consider the term containing the one electron operators  $G(j)$ , which we denote by  $E_1$ . It is easily seen that

$$E_1 = 2 \sum_j G_j = 2 \sum_j \langle \phi_j | G | \phi_j \rangle \quad (11-12)$$

since any permutation of the function  $\Psi_0$  on the right side of Eq. (1-11) will be zero because of the orthogonality condition (11-9) of the orbitals  $\phi_i$ .

The contribution of the two-electron operators  $\Omega(j, k)$  to  $E$  may be separated into two parts  $E_2$  and  $E_3$ . The part  $E_2$  is the contribution of the non-permuted term on the right side of Eq. (11-11). It is given by

$$E_2 = \sum_{i=1}^N J_{i,i} + 4 \sum_{j>i} J_{i,j} \quad (11-13)$$

Here the integrals  $J_{i,j}$  are known as Coulomb integral and they are defined as

$$J_{i,j} = \langle \phi_i(1)\phi_j(2) | \Omega(1,2) | \phi_i(1)\phi_j(2) \rangle \quad (11-14)$$

It will prove to be convenient to rewrite Eq. (11-14) as

$$\begin{aligned} E_2 &= \sum_{i=1}^N J_{i,i} + 2 \sum_{i=1}^N \sum_{j \neq i} J_{i,j} \\ &= 2 \sum_{i=1}^N \sum_{j=1}^N J_{i,j} - \sum_{i=1}^N J_{i,i} \end{aligned} \quad (11-15)$$

The second part  $E_3$  of the contributions of the operators  $\Omega(j, k)$  to  $E$  is due to all possible permutations in Eq. (11-11) but we should realize that only a fraction of these permutations lead to a nonzero result. Because of the orthogonality of the spin functions we only obtain a nonzero result if we permute either within the set of even or within the set of odd numbered electrons. Subject to this restraint we may only exchange one pair of electrons in order to get a nonzero result because of the orthogonality (11-9) of the orbitals. The result is

$$E_3 = -2 \sum_{j>i} K_{i,j} \quad (11-16)$$

The exchange integrals  $K_{i,j}$  are here defined as

$$K_{i,j} = \langle \phi_i(1)\phi_j(2) | \Omega(1,2) | \phi_j(1)\phi_i(2) \rangle \quad (11-17)$$

We may again rearrange Eq. (11-16) as

$$E_3 = - \sum_{i=1}^N \sum_{j=1}^N K_{i,j} + \sum_{i=1}^N K_{i,i} \quad (11-18)$$

The desired expression for the expectation value  $E$  is now obtained by taking the sum of the three contributions (11-12), (11-15) and (11-18) which gives

$$\begin{aligned} E &= 2 \sum_i G_i + 2 \sum_i \sum_j J_{i,j} - \sum_i \sum_j K_{i,j} \\ &= \sum_{i=1}^N \left[ 2G_i + \sum_{j=1}^N (2J_{i,j} - K_{i,j}) \right] \end{aligned} \quad (11-19)$$

since

$$J_{i,i} = K_{i,i} \quad (11-20)$$

The Hartree-Fock equations may now be derived from the energy expression (11-19) by varying one of the orbitals, for example  $\phi_k$ , by an amount  $\delta\phi_k$  and by setting the corresponding change  $\delta_k E$  in the energy equal to zero,

$$\delta_k E = 0 \quad (11-21)$$

It is allowed to vary only the functions  $\phi_k^*$  on the left of the operators without loss of generality and it follows then that

$$\begin{aligned} \delta_k E &= 2 \langle \delta\phi_k | G | \phi_k \rangle + 4 \sum_{j=1}^N \langle \delta\phi_k(1)\phi_j(2) | \Omega(1,2) | \phi_k(1)\phi_j(2) \rangle \\ &\quad - 2 \sum_{j=1}^N \langle \delta\phi_k(1)\phi_j(2) | \Omega(1,2) | \phi_j(1)\phi_k(2) \rangle \end{aligned} \quad (11-22)$$

This expression may be simplified by introducing the operator

$$J_j(\mathbf{r}_1) = \int \phi_j^*(\mathbf{r}_2) \Omega(\mathbf{r}_1, \mathbf{r}_2) \phi_j(\mathbf{r}_2) d\mathbf{r}_2 \quad (11-23)$$

It is possible to transform the exchange integrals in a similar fashion but in the latter case the definition of the operator becomes more complex. We define the operators  $K_j(\mathbf{r}_1)$  by means of the equation

$$K_j(\mathbf{r}_1)\psi(\mathbf{r}_1) = \left\{ \int \phi_j^*(\mathbf{r}_2) \Omega(\mathbf{r}_1, \mathbf{r}_2) \psi(\mathbf{r}_2) d\mathbf{r}_2 \right\} \phi_j(\mathbf{r}_1) \quad (11-24)$$

It is easily verified that the Coulomb integral in Eq. (11-22) may now be written as

$$\langle \delta\phi_k(1)\phi_j(2) | \Omega(1,2) | \phi_k(1)\phi_j(2) \rangle = \langle \delta\phi_k(1) | J_j(1) | \phi_k(1) \rangle \quad (11-25)$$

while the exchange type integrals may be represented as

$$\langle \delta\phi_k(1)\phi_j(2)|\Omega(1,2)|\phi_j(1)\phi_k(2)\rangle = \langle \delta\phi_k(1)|K_j(1)\phi_k(1)\rangle \quad (11-26)$$

By substituting these results into Eq. (11-22) we obtain

$$\delta_k E = 2\langle \delta\phi_k|G|\phi_k\rangle + 4\sum_{j=1}^N \langle \delta\phi_k|J_j|\phi_k\rangle - 2\sum_{j=1}^N \langle \delta\phi_k|K_j|\phi_k\rangle \quad (11-27)$$

By introducing the Hartree-Fock operator

$$F_{\text{op}} = G + \sum_{j=1}^N (2J_j - K_j) \quad (11-28)$$

We may reduce Eq (11-27) to the simple form

$$\delta_k E = 2\langle \delta\phi_k|F_{\text{op}}|\phi_k\rangle = 0 \quad (11-29)$$

subject to the restraint

$$\langle \delta\phi_k|\phi_k\rangle = 0 \quad (11-30)$$

It is worth noting that the Hartree-Fock operator contains a sum over all occupied orbitals  $\phi_j$  so that all occupied orbitals contribute equally to the operator. Even though one of the orbitals  $\phi_k$  was selected as the orbital to be varied this orbital does not have a preferred role in the definition of the Hartree-Fock operator. Consequently it follows from Eqs. (11-28) and (11-29) that all occupied orbitals must be eigenfunctions of the same Hartree-Fock operator  $F_{\text{op}}$ . These eigenfunctions  $\phi_k$  and corresponding eigenvalues  $\lambda_k$  are defined as

$$F_{\text{op}}\phi_k = \lambda_k\phi_k \quad (11-31)$$

We should realize that the Hartree-Fock operator  $F_{\text{op}}$  is constructed from a set of approximate orbitals. Even the exact solutions of the eigenvalue problem (11-31) are therefore of an approximate nature. On the other hand, it may be assumed that the set of orbitals which are the solutions of Eq. (11-31) are more accurate than the set of orbitals that were used in constructing the operator  $F_{\text{op}}$ . It is therefore advantageous to define an improved operator  $F_{\text{op}}$  by substituting the solutions of the previous operator. The solutions of the new and improved operator  $F_{\text{op}}$  should be more accurate than the solutions of the previous operator. In the SCF method this procedure is repeated a number of times until the solutions of the eigenvalue



problem (11-31) become identical with the set of orbitals that were used in constructing the Hartree-Fock operator. It is said that at this point self-consistency is achieved, hence the name Self Consistent Field or SCF method. It is also customary to refer to the results of the SCF procedure as the solutions of the Hartree-Fock method. It appears that either one of the two names, SCF or HF, is generally accepted for the description of the method.

It may be derived from the definition (11-28) of the Hartree Fock operator  $F_{\text{op}}$  that the eigenvalue  $\lambda_k$  is given by

$$\lambda_k = \langle \phi_k | F_{\text{op}} | \phi_k \rangle = G_k + \sum_{j=1}^N (2J_{k,j} - K_{k,j}) \quad (11-32)$$

It is important to note that the total energy as defined by Eq. (11-19) is not equal to the sum of the Hartree-Fock parameters but it is instead given by

$$E = \sum_{k=1}^N (G_k + \lambda_k) \quad (11-33)$$

The Hartree-Fock eigenvalue problem (11-31) has in principle an infinite number of eigenvalues and corresponding eigenfunctions. If we assume that self-consistency has been achieved, then the set of eigenfunctions  $\phi_1, \phi_2, \dots, \phi_N$  corresponding to the lowest  $N$  eigenvalues represents the filled orbitals of the system. It may be argued that the additional eigenfunctions  $\phi_{N+1}, \phi_{N+2}$ , and so on, corresponding to higher eigenvalues  $\lambda_{N+1}, \lambda_{N+2}$ , and so on, have no physical meaning. However, it is generally assumed that these additional Hartree-Fock eigenfunctions may be used for the construction of the wave functions corresponding to either singlet or triplet excited molecular configurations. The specific form of these excited state wave functions was presented in Eqs. (11-5) and (11-6). The corresponding excitation energies are given by

$$\begin{aligned} {}^1E(j \rightarrow n) - E &= \langle {}^1\Psi \langle j \rightarrow n | H - E | {}^1\Psi(j \rightarrow n) \rangle \\ &= \lambda_n - \lambda_j - J_{j,n} + 2K_{j,n} \end{aligned} \quad (11-34)$$

and

$$\begin{aligned} {}^3E(j \rightarrow n) - E &= \langle {}^3\Psi(j \rightarrow n) | H - E | {}^3\Psi(j \rightarrow n) \rangle \\ &= \lambda_n - \lambda_j - J_{j,n} \end{aligned} \quad (11-35)$$

where the energy  $E$  is defined in Eq. (11-8).

There is a much simpler relation between the Hartree-Fock parameters  $\lambda_k$  and the ionization energies of the system. The energy required to remove an electron from a doubly occupied orbital  $\phi_k$  is approximately equal to the corresponding eigenvalue  $\lambda_k$ . This theorem was proved in 1934 by *Tjalling Charles Koopmans*

(1910–1985), who was a graduate student in theoretical physics with Kramers at the time. It is interesting to note that this is the only contribution to theoretical physics by Koopmans because his interest changed to mathematical economics. He received the 1975 Nobel Prize in economics.

#### IV. SLATER ORBITALS

It is necessary to know a set of approximate atomic orbitals as initial input for the SCF method, and it is, of course, also useful to be able to predict the approximate form of the atomic Hartree-Fock orbitals. Such a set of orbitals may be derived by means of arguments that are similar to those in our discussion of the helium atom in Section 10.IV. There we considered the  $(1s)^2$  ground state configuration, and we argued that the  $1s$  atomic orbital could be approximated by the function

$$\phi(1s) = (q^3/\pi)^{1/2} \exp(-qr) \quad (11-36)$$

By making use of the variational principle, we found that the lowest energy is obtained if  $q = 1.6875$ . It is possible to derive similar approximate orbitals for more complex atoms by means of similar applications of the variational principle.

It may be helpful to recall the physical reasoning that we used in proposing the above form (11-36) for the atomic orbital. We argued that each of the two electrons experiences a force that is the difference between the attractive force of the positively charged nucleus and the repulsive force of the negatively charged second electron. The effect of the repulsion by the second electron can be roughly represented as a shielding of the nuclear charge  $2e$  by an amount  $\sigma e$  due to the fraction of the charge cloud of the second electron that is situated between the nucleus and the first electron. The value of  $\sigma$  should obviously be somewhere between 0 and 1, and the application of the variation principle gives the result

$$\sigma = 2 - q = 0.3125 \quad (11-37)$$

The physical arguments that were applied to the helium atom may also be extended to more complex systems. For example, in the case of the neon atom, which has the electronic configuration  $(1s)^2 (2s)^2 (2p)^6$  we introduce the following set of approximate orbitals:

$$\begin{aligned} \phi(1s) &= (q_1^3/\pi)^{1/2} \exp(-q_1 r) \\ \phi(2s) &= (q_2^3/2\pi)^{1/2} (q_2 r - 1) \exp(-q_2 r) \\ \phi(2p) &= (q_2^5/\pi)^{1/2} r_\alpha \exp(-q_3 r) \quad \alpha = x, y, z \end{aligned} \quad (11-38)$$

The values of the orbital exponents  $q_1$ ,  $q_2$ , and  $q_3$  can then be determined by means of the variational principle.

A calculation of the above type was performed for the carbon atom. In this case, it was found that  $q_1$  is very close to the nuclear charge  $Z = 6$ , while the orbital exponents  $q_2$  and  $q_3$  for the  $2s$  and  $2p$  orbitals are both slightly larger than 1.5. The difference between  $q_2$  and  $q_3$  was found to be quite small. It should be noted that the two orbitals  $\phi(1s)$  and  $\phi(2s)$  are no longer orthogonal if  $q_2$  is different from the value  $(q_1/2)$ . It is therefore more convenient to introduce a different approximate function

$$\phi(2s) = (q_2^5/3\pi)^{1/2} r \exp(-q_2 r) \quad (11-39)$$

for the  $2s$  orbital and base the calculation on an orthogonalized version:

$$\begin{aligned} \phi'(2s) &= a\phi(2s) - b\phi(1s) \\ \langle \phi'(2s) | \phi(1s) \rangle &= 0 \quad \langle \phi'(2s) | \phi'(2s) \rangle = 1 \end{aligned} \quad (11-40)$$

In 1930 Slater proposed a set of simple algebraic rules for estimating the values of the effective nuclear charges and corresponding orbital exponents in atomic orbitals of the type shown in Eqs. (11-38) and (11-39). Slater's empirical rules were based on the available information that had been derived from variational calculations. The resulting orbitals became widely known as Slater orbitals, and they are now described by the abbreviation STO (Slater-type orbitals).

The (unnormalized) Slater orbitals are given by

$$\begin{aligned} \phi(ns) &= r^{n-1} \exp(-qr) & q &= Z_{\text{eff}}/n \\ \phi(np_\alpha) &= r^\alpha r^{n-2} \exp(-qr) & q &= Z_{\text{eff}}/n \\ \alpha &= x, y, z & & \text{etc.} \end{aligned} \quad (11-41)$$

The effective nuclear charges  $Z_{\text{eff}}$  are described by the equation

$$Z_{\text{eff}} = Z - \sigma \quad (11-42)$$

where  $Z$  is the exact nuclear charge and the parameter  $\sigma$  represents the shielding of the nucleus by the other electrons.

The value of  $\sigma$  depends on the state of the electron that we are concerned with and on the states of the other electrons present in the atom. According to Slater's rules,  $\sigma$  is obtained as a sum of the shielding contributions of these other electrons. These contributions are:

1. Nothing from any electron that has a principal quantum number  $n$  that is higher than the one we consider.
2. An amount 0.35 from each electron that has the same principal quantum number as the electron that we consider, except that when we consider a  $(1s)$  electron, the contribution from the other  $(1s)$  electron is 0.30.

3. An amount 0.85 from each electron that has a principal quantum number  $n$  that is one less than the quantum number of the electron that we consider if the latter is an  $s$  or a  $p$  electron, and an amount 1.00 from each electron whose principal quantum number is one less than the electron that we consider if the latter is in a  $d$ ,  $f$ , or  $g$  state.
4. An amount 1.00 from each electron with a principal quantum number that is less by two or more than the quantum number of the electron considered.

We illustrate the application of these rules to a few selected atoms, namely, helium, carbon, and sulfur. In the case of the helium atom, the Slater rules predict that the effective nuclear charge is given by

$$Z_{\text{eff}} = 2 - 0.30 = 1.70 \quad (11-43)$$

which agrees reasonably well with the result of the variational treatment. The electronic configuration of the carbon atom is  $(1s)^2 (2s)^2 (2p)^2$  and the Slater rules predict the following values of the effective nuclear charges:

$$\begin{aligned} Z_{\text{eff}}(1s) &= 6 - 0.30 = 5.70 \\ Z_{\text{eff}}(2s) &= Z_{\text{eff}}(2p) = 6 - 3 \times 0.35 - 2 \times 0.85 = 3.25 \end{aligned} \quad (11-44)$$

A previous variational calculation had predicted a value of 3.18 for  $Z_{\text{eff}}(2p)$ , which is not too different from the corresponding Slater value. In the case of the sulfur atom, the electronic configuration is  $(1s)^2 (2s)^2 (2p)^6 (3s)^2 (3p)^4$  and the Slater effective nuclear charges are

$$\begin{aligned} Z_{\text{eff}}(1s) &= 16 - 0.30 = 15.70 \\ Z_{\text{eff}}(2s) &= Z_{\text{eff}}(2p) = 16 - 7 \times 0.35 - 2 \times 0.85 = 11.85 \\ Z_{\text{eff}}(3s) &= Z_{\text{eff}}(3p) = 16 - 2 \times 1.00 - 8 \times 0.85 - 5 \times 0.35 = 5.45 \end{aligned} \quad (11-45)$$

The Slater-type orbitals have been used in computer programs for molecular structure calculations such as the Gaussian Program packages, but they always corresponded to the lowest level of approximation in the calculations, and their use is no longer considered acceptable.

## V. MULTIPLET THEORY

In Section 11.III we derived the Hartree-Fock equations for a closed-shell nondegenerate ground state, but we should realize that many atomic electronic configurations have a high degree of degeneracy in a first approximation. For instance, an atomic  $p$  orbital is threefold degenerate and the corresponding spin is twofold

degenerate, so that the configuration  $(1s)^2 (2s)^2 (2p)^2$  of the carbon atom corresponds to 36 different eigenstates to a first approximation. A more detailed analysis that takes into account the small interactions between the orbital and spin magnetic moments of the  $(2p)$  electrons will predict that the 36-fold degenerate eigenstate will be resolved into a number of eigenstates with lower degrees of degeneracy. Such a precise analysis would require the diagonalization of a  $36 \times 36$  matrix, which is a fairly laborious task. Fortunately, there is a much simpler approach to the description of atomic eigenstates known as the vector model of atomic structure. The vector model is based on general considerations, and it leads to a satisfactory qualitative interpretation of atomic spectra even though it may not yield exact quantitative predictions. There are various types of vector models, but we discuss only the most common type, known as Russell-Saunders coupling, which describes situations where the interactions between the orbital and spin angular momenta may be assumed to be small. This assumption is valid in the majority of cases.

The complete atomic Hamiltonian  $H$  is the sum of three terms:

$$H = H_o + H_{so} + H_{ss} \quad (11-46)$$

The first term  $H_o$  is the spinless atomic Hamiltonian that we have considered up to now. The second term  $H_{so}$  is known as the spin-orbit coupling. It represents the interaction between the orbital and the spin magnetic moments. There is no need to consider the third term  $H_{ss}$ , the spin-spin coupling, which represents interactions between the various electron spins.

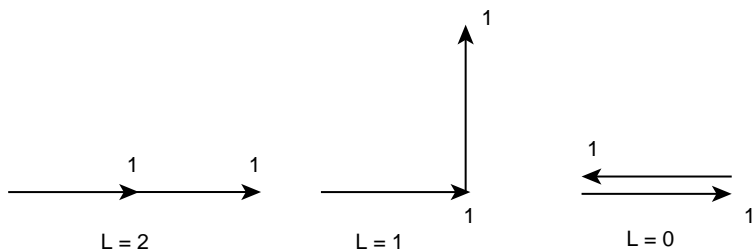
We showed in Chapter 7 that the eigenstates of the Hamiltonian  $H_o$  are all characterized by the values of the total orbital angular momentum  $\mathbf{L}$ , which is the sum of the orbital angular momenta  $\mathbf{l}_i$  of the individual electrons:

$$\mathbf{L} = \sum_i \mathbf{l}_i \quad (11-47)$$

This is due to the fact that the components of  $\mathbf{L}$  and its magnitude  $\mathbf{L}^2$  all commute with  $H_o$ . In a similar fashion, we mentioned in Chapter 10 that the eigenstates of  $H_o$  are also characterized by the value of the total spin angular momentum  $\mathbf{S}$ , which is the sum of the spin angular momenta  $\mathbf{s}_i$  of the individual electrons:

$$\mathbf{S} = \sum_i \mathbf{s}_i \quad (11-48)$$

Within a given electronic configuration, the atomic energies depend on  $S$  and to a lesser extent on  $L$ . According to Hund's rule mentioned in Section 11.I, the state with the lowest energy is the state with the largest value of  $S$  and, for a given  $S$  value, the state with the largest value of  $L$ .



**Figure 11-1** Addition of two orbital angular momentum vectors of magnitude 1.

The total atomic angular momentum  $\mathbf{J}$  is the vector sum of the orbital angular momentum  $\mathbf{L}$  and the spin angular momentum  $\mathbf{S}$ ;

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \quad (11-49)$$

$$J^2 = L^2 + S^2 + 2(\mathbf{L} \cdot \mathbf{S}) \quad (11-50)$$

The Russell-Saunders coupling scheme is based on the assumption that the spin-orbit coupling is much smaller than the coupling between the orbital angular momenta and between the spin angular momenta of the individual electrons. In that case, we may first determine the values of  $L$  and  $S$ . The possible values of  $L$  are obtained by quantizing the orbital angular momentum of each electron relative to that of the other electrons.

As an example, we will consider the  $(2p)(3p)$  configuration. We quantize the  $3p$  angular momentum relative to the  $2p$  angular momentum (see Figure 11-1), and we find that

$$L = 2, 1, 0 \quad (11-51)$$

We follow the same procedure for the spin angular momenta, (see Figure 10-1), and we find that

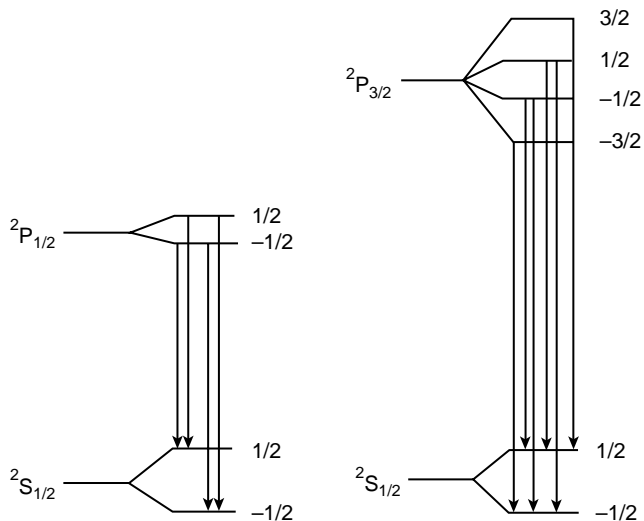
$$S = 1, 0 \quad (11-52)$$

The six possible eigenstates are now

$${}^1S_0, {}^1P_1, {}^1D_2, {}^3S_1, {}^3P_2, {}^3D_3 \quad (11-53)$$

The possible values of  $J$  are finally obtained by quantizing the smaller of the two vectors relative to the larger one. This leads to the following states:

$${}^1S_0, {}^1P_1, {}^1D_2, {}^3S_1, {}^3P_2, {}^3P_1, {}^3P_0, {}^3D_3, {}^3D_2, {}^3D_1 \quad (11-54)$$



**Figure 11-2** The anomalous Zeeman effect of the two sodium D lines. In a magnetic field the  $D_1$  line splits into four lines and the  $D_2$  line splits into six lines.

If we consider the configuration  $(2p)^2$ , then the exclusion principle must be taken into account, and it is found that only three eigenstates— $^1S$ ,  $^1D$ , and  $^3P$ —are allowed.

In the Russell-Saunders coupling scheme, the spin-orbit interaction is treated as a small perturbation. Its net effect on the energy levels is the splitting of the degenerate energy levels with given values of the quantum numbers  $L$  and  $S$ . The major part of the spin-orbit coupling transforms as the scalar product  $(\mathbf{L} \cdot \mathbf{S})$ . The magnitude of the spin-orbit coupling depends therefore on the relative orientation of the two vectors  $\mathbf{L}$  and  $\mathbf{S}$ , which means on the value of the quantum number  $J$  according to Eq. (11-50).

As an illustration, we will explain the splitting of the sodium D line (see Figure 11-2). It can be seen in Table 11-1 that the ground-state configuration of the sodium atom is  $(1s)^2 (2s)^2 (2p)^6 (3s)$ . In this configuration  $L = 0$  and  $S = 1/2$ . Consequently,  $J = 1/2$  and the ground state is described by the symbol  $^2S_{1/2}$ , as indicated in the table. The sodium D line corresponds to a transition from the ground state to the states belonging to the configuration  $(1s)^2 (2s)^2 (2p)^6 (3p)$ . The latter has the quantum numbers  $L = 1$  and  $S = 1/2$ . There are therefore two possible values of the quantum numbers  $J$ , namely,  $J = 1/2$  and  $J = 3/2$ , and two different atomic states,  $^2P_{1/2}$  and  $^2P_{3/2}$ , belonging to the configuration  $(1s)^2 (2s)^2 (2p)^6 (3p)$ .

The two  $^2P$  states have slightly different energies due to spin-orbit perturbation, and consequently there are two spectral lines corresponding to the  $^2S \rightarrow ^2P$  transition. The sodium  $D_1$  line at  $5895.93 \text{ \AA}$  corresponds to the  $^2S_{1/2} \rightarrow ^2P_{1/2}$  transition, and the sodium  $D_2$  line at  $5889.96 \text{ \AA}$  corresponds to the  $^2S_{1/2} \rightarrow ^2P_{3/2}$  transition.

The difference in energy between the two  ${}^2P$  energy levels is  $17.19 \text{ cm}^{-1}$  in magnitude, and it is due to the spin-orbit perturbation.

Figure 11-2 also shows that the energy levels split into  $(2J + 1)$  components in the presence of a magnetic field; consequently, the sodium  $D_1$  line splits into four components and the sodium  $D_2$  line splits into six components. This is the explanation of the anomalous Zeeman effect mentioned in Section 10.II.

## VI. CONCLUDING REMARKS

In our discussion of atomic structure, we have emphasized those aspects that are useful for an understanding the quantum theory of molecules. Since the majority of molecular wave functions are expanded in terms of atomic orbitals of one type or another, it is useful to be familiar with the various atomic orbitals that have been derived.

We also believed that it would be useful to present a detailed derivation of the Hartree-Fock equations since this method is the basis of most molecular computational programs. The difference between atomic and molecular theories is that the atomic Hartree-Fock equations can be solved exactly, if only in numerical form, while the solution of the molecular Hartree-Fock equations requires additional approximations. We will address these and other issues in the following chapter.

## VII. PROBLEMS

- 11-1 Write the antisymmetrized wave function including spin of the  $(1s)^2 (2s)$  configuration of the Li atom expressed in terms of the orthonormal orbitals  $s_1$  and  $s_2$ .
- 11-2 Write the antisymmetrized singlet and triplet wave functions including spin of the excited  $(1s)^2 (2s) (3s)$  configuration of the Be atom expressed in terms of the orthonormal orbitals  $s_1$ ,  $s_2$  and  $s_3$ .
- 11-3 Derive expressions for the expectation values of the Hamiltonian of the  $(1s)^2 (2s)^2$  and of the singlet and triplet configurations  $(1s)^2 (2s) (3s)$ , all expressed in terms of the orthonormal orbitals  $s_1$ ,  $s_2$ ,  $s_3$ . Derive also the singlet and triple excitation energies.
- 11-4 Derive the energy difference between the  $(1s)^2 (2s)^2 (2p)^6 (3s)$  configuration of the Na atom and the  $(1s)^2 (2s)^2 (2p)^6$  configuration of the  $\text{Na}^+$  ion expressed in terms of atomic orbitals. Compare the result with the Hartree-Fock eigenvalues  $\lambda_k$ .
- 11-5 Derive the Slater orbitals of the phosphorus atom.
- 11-6 Derive the Slater orbitals of the sodium, potassium, and calcium atoms.



- 11-7** Calculate the expectation values of the coordinate  $r$  for the Slater orbitals of the valence electrons of the sodium, potassium and calcium atoms. Compare the three results and interpret their relative magnitudes.
- 11-8** What are the possible values of the orbital angular momentum in the  $(1s)^2 (2s) (2p)^3$  configuration of the carbon atom?
- 11-9** Which of the possible eigenstates of the  $(1s)^2 (2s) (2p)^3$  configuration of the carbon atom has the lowest energy according to Hund's rule?
- 11-10** List all possible eigenstates (and their degeneracies) of the configuration  $(p) (p')$  that are derived before spin-orbit coupling is taken into account.

---

# 12

---

## MOLECULAR STRUCTURE

### I. INTRODUCTION

In 1929 Dirac expressed the opinion that the fundamental principles of quantum theory were well established. We quote a few key sentences of his famous statement: “The general theory of quantum mechanics is now almost complete, the imperfections that still remain being in connection with the exact fitting in of the theory with relativity ideas. These give rise to difficulties only when high-speed particles are involved, and are therefore of no importance in the consideration of atomic and molecular structure and ordinary chemical reactions. . . . The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.”

It may well be true that the above-mentioned equations were too complicated to be solved exactly, but some ambitious scientists felt that it should be possible to derive approximate solutions. This led to the establishment of a new scientific discipline, quantum chemistry, which is concerned with the application of quantum theory to the elucidation of molecular structure and to the prediction of molecular properties.

The major problem in quantum chemistry has always been the determination of the molecular eigenfunctions. Exact analytical solutions of the Schrödinger equation have been derived only for the hydrogen atom and, to some extent, for the

hydrogen molecular ion. The eigenfunctions of more complex systems may only be obtained in approximate form by means of approximate methods.

Quite soon, two different approaches began to emerge in quantum chemistry. In 1960 *Charles Alfred Coulson* (1910–1974) proposed in an after-dinner speech that the quantum chemists could be separated into two very distinct groups, which he called the ab-initio-ists and the a-posterio-ists. The first group was interested in deriving highly accurate molecular eigenfunctions, and their efforts were usually confined to small molecules such as diatomics or even the hydrogen molecule. The second group focused on larger systems, frequently aromatic or conjugated organic molecules. The wave functions that they used in their calculations were often obtained by educated guesses rather than from mathematical derivations. Their goal was to make reliable predictions about the properties of large molecules even if those predictions lacked a sound mathematical foundation.

In general, the ab-initio-ists were often trained as physicists or mathematicians, and they were primarily interested in deriving better mathematical methods; many of their early efforts dealt with the hydrogen molecule. The a-posterio-ists were often guided by chemical intuition rather than by mathematics, and many of them were trained as chemists.

Eventually the two approaches merged because the more precise procedures that were developed for small molecules could be applied to larger and larger systems, especially after high-speed computers became available. The successful developments in quantum chemistry were officially recognized by the award of the 1998 Nobel Prize in chemistry to *John Anthony Pople* (1924–) and *Walter Kohn* (1923–).

In this chapter we present a brief overview of the application of quantum theory to molecular structure. We first describe the separation of nuclear and electronic motion, illustrated for diatomic molecules. Next, we discuss the nature of the chemical bond by taking the hydrogen molecule as an example and by describing the approximate electronic structures of some simple diatomic and polyatomic molecules. Finally, we explain the approximate theories that were developed for aromatic and conjugated molecules, in particular the approach of *Erich Armand Arthur Joseph Hückel* (1896–1980).

## II. THE BORN-OPPENHEIMER APPROXIMATION

A molecule contains both nuclei and electrons, but the nuclear and electronic motions may be considered separately because the nuclei are much heavier than the electrons and their motion is therefore more restricted. The separability of nuclear and electronic motion is known as the Born-Oppenheimer approximation.

*Jules Robert Oppenheimer* (1904–1967) was a young American who went to Europe to study physics after graduating from Harvard. He worked with Born in Göttingen, where he was awarded a Ph.D. degree in 1927. The separability of nuclear and electronic motion was formulated in 1927 by Born and Oppenheimer

by means of a precise but rather complex mathematical analysis. We present a less comprehensive but simpler version of their work.

We first describe a quantum mechanics theorem that deals with the separability of the Hamiltonian. We consider a Hamiltonian  $H(X, Y)$  that depends on two sets of coordinates,  $X$  and  $Y$ , and that may be written as a sum of two parts:

$$H(X, Y) = H_1(X) + H_2(Y) \quad (12-1)$$

Here the first part,  $H_1$  depends only on the coordinates  $X$  and the second part,  $H_2$ , depends only on the coordinates  $Y$ .

We define the eigenvalues of  $H_1$  and  $H_2$  as

$$\begin{aligned} H_1(X)\Phi_n(X) &= \lambda_n\Phi_n(X) \\ H_2(Y)\Psi_m(Y) &= \mu_m\Psi_m(Y) \end{aligned} \quad (12-2)$$

It is then easily verified that

$$H(X, Y)\Phi_n(X)\Psi_m(Y) = (\lambda_n + \mu_m)\Phi_n(X)\Psi_m(Y) \quad (12-3)$$

We see that the eigenfunctions of the operator  $H(X, Y)$  of Eq. (12-1) are products of the eigenfunctions of  $H_1$  and  $H_2$ , while its eigenvalues are sums of the eigenvalues of  $H_1$  and  $H_2$ .

A diatomic molecule contains two nuclei, a and b, with masses  $M_a$  and  $M_b$ , electric charges  $Z_a e$  and  $Z_b e$ , and coordinates  $\mathbf{R}_a$  and  $\mathbf{R}_b$ . In addition, there are  $N$  electrons with coordinates  $\mathbf{r}_i$ . The molecular Hamiltonian may now be written as a sum of a nuclear and an electronic part,

$$H_{\text{mol}} = H_{\text{nucl}} + H_{\text{el}} \quad (12-4)$$

We write the nuclear part as

$$H_{\text{nucl}} = -\frac{\hbar^2}{2M_a}\Delta_a - \frac{\hbar^2}{2M_b}\Delta_b + \frac{Z_a Z_b e^2}{R_{ab}} \quad (12-5)$$

Here  $\Delta_a$  and  $\Delta_b$  are Laplace operators

$$\Delta_a = \frac{\partial^2}{\partial X_a^2} + \frac{\partial^2}{\partial Y_a^2} + \frac{\partial^2}{\partial Z_a^2} \quad \text{etc.} \quad (12-6)$$

and  $R_{ab}$  is the internuclear distance. The electronic part is

$$H_{\text{el}} = -\sum_j \left( \frac{\hbar^2}{2m}\Delta_j + \frac{e^2}{r_{aj}} + \frac{e^2}{r_{bj}} \right) + \sum_{j>i} \frac{e^2}{r_{ij}} \quad (12-7)$$

In order to analyze the eigenfunctions of the molecular Hamiltonian, we replace the nuclear coordinates  $\mathbf{R}_a$  and  $\mathbf{R}_b$  by the coordinates  $\mathbf{R}_c$  of their center of gravity and  $\mathbf{R}$  of their distance:

$$\mathbf{R}_c = \frac{M_a \mathbf{R}_a + M_b \mathbf{R}_b}{M_a + M_b} \quad \mathbf{R} = \mathbf{R}_b - \mathbf{R}_a \quad (12-8)$$

It is easily shown that the motion of the center of gravity may then be separated; it may therefore be disregarded. The nuclear Hamiltonian is then reduced to

$$H_{\text{nuc1}} = -\frac{\hbar^2}{2\mu} \left( \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2} \right) + \frac{Z_a Z_b e^2}{R} \quad (12-9)$$

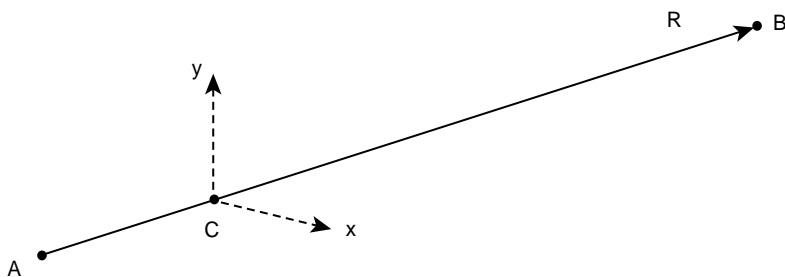
where  $\mu$  is the reduced mass of the nuclei a and b:

$$\frac{1}{\mu} = \frac{1}{M_a} + \frac{1}{M_b} \quad (12-10)$$

We now define the nuclear center of gravity as the origin of the electron coordinates and the vector  $\mathbf{R}$  as their  $Z$  axis (see Figure 12-1). We note that the electronic Hamiltonian  $H_{\text{el}}$  of Eq. (12-7) contains the electron coordinates that we denote by  $r$ , but it also depends implicitly on the internuclear distance  $R$ . We write it therefore as  $H_{\text{el}}(r; R)$ . The total molecular Hamiltonian of Eq. (12-4) may now be represented as

$$H_{\text{mol}}(r; R) = H_{\text{nuc1}}(R) + H_{\text{el}}(r; R) \quad (12-11)$$

The separation of nuclear and electronic motion is not immediately obvious because Eq. (12-11) differs from Eq. (12-1) insofar as  $H_{\text{el}}$  depends on the nuclear coordinate  $R$  in addition to the electronic coordinates  $r$ . However, the two types of motion may still be separated if we make the two assumptions that are equivalent with the Born-Oppenheimer approximation.



**Figure 12-1** Definition of nuclear and electronic coordinate systems in a diatomic molecule.

We first define the eigenvalues  $E_n$  and corresponding eigenfunctions  $F_n$  of the operator  $H_{\text{el}}(r; R)$  by means of

$$H_{\text{el}}(r; R)F_n(r; R) = \varepsilon_n(R)F_n(r; R) \quad (12-12)$$

We note that both the eigenvalues and eigenfunctions depend on  $R$ .

Our first assumption is that the eigenfunctions of the molecular Hamiltonian  $H_{\text{mol}}$  may be represented as products of a function  $f_n(X, Y, Z)$  of the nuclear coordinates only and of one of the eigenfunctions  $F_n(r; R)$

$$H_{\text{mol}}(r; R)f_n(X, Y, Z)F_n(r; R) = E_n f_n(X, Y, Z)F_n(r; R) \quad (12-13)$$

Our second assumption is

$$\frac{\partial^2}{\partial X^2} [f_n(X, Y, Z)F_n(r; R)] = F_n(r; R) \frac{\partial^2 f_n(X, Y, Z)}{\partial X^2}, \text{ etc.} \quad (12-14)$$

In other words, the derivatives of the electronic eigenfunction  $F_n$  with respect to the nuclear coordinates  $(X, Y, Z)$  are much smaller than the corresponding derivatives of the nuclear functions  $f_n$ . It may be helpful to present a simple physical interpretation of the Born-Oppenheimer approximation. First, we may derive the electronic part of the electronic wave function on the assumption that the nuclei are stationary at a fixed internuclear distance  $R$ . Second, even though the electronic wave function is dependent on the internuclear distance  $R$ , its changes as a function of  $R$  are negligible compared to the changes in the nuclear wave function. In summary, even though the nuclear and electronic motions may not be separated in the strictest sense, for all practical purposes it is permitted to separate them.

### III. NUCLEAR MOTION OF DIATOMIC MOLECULES

It has been well established from spectroscopic measurements that the total molecular energy of a diatomic molecule is the sum of an electronic, a vibrational, and a rotational energy:

$$E(n, \nu, J) = \varepsilon_n + \left( \nu + \frac{1}{2} \right) \hbar \omega_n + J(J+1)B_{n,\nu} \quad (12-15)$$

Here  $n$ ,  $\nu$ , and  $J$  are the electronic, vibrational, and rotational quantum numbers, respectively;  $\omega_n$  is the vibrational frequency corresponding to the electronic state  $n$ ; and  $B_{n,\nu}$  is its rotational constant. We will derive the above relation by solving the molecular Schrödinger equation

$$[H_{\text{nuc}}(R) + H_{\text{el}}(r; R)]\Psi(r; R) = E\Psi(r; R) \quad (12-16)$$

while making use of the Born-Oppenheimer approximation.

We assume that the electronic eigenstate of the molecule is described by Eq. (12-12) so that we may substitute

$$\Psi(r; R) = f_n(X, Y, Z)F_n(r; R) \quad (12-17)$$

into the Schrödinger equation (12-16). The result is

$$[H_{\text{nucl}}(R) + \varepsilon_n(R)]f_n(X, Y, Z) = Ef_n(X, Y, Z) \quad (12-18)$$

since we may divide the Schrödinger equation by the electronic eigenfunction  $F_n(r; R)$ .

This equation may be solved by transforming the nuclear coordinates  $(X, Y, Z)$  into polar coordinates  $(R, \theta, \phi)$ ,

$$\begin{aligned} X &= R \sin \theta \cos \phi \\ Y &= R \sin \theta \sin \phi \\ Z &= R \cos \theta \end{aligned} \quad (12-19)$$

and by introducing the nuclear angular momentum operators

$$L_x = \frac{\hbar}{i} \left( Y \frac{\partial}{\partial Z} - Z \frac{\partial}{\partial Y} \right) \text{ etc.} \quad (12-20)$$

analogous to Eq. (7-6). This allows us to write the nuclear Hamiltonian  $H_{\text{nucl}}$  of Eq. (12-9) in the form

$$H_{\text{nucl}} = -\frac{\hbar^2}{2\mu} \left( \frac{\partial^2}{\partial R^2} + \frac{2}{R} \frac{\partial}{\partial R} \right) + \frac{L^2}{2\mu R^2} + \frac{Z_a Z_b e^2}{R} \quad (12-21)$$

Here we have made use of Eqs. (3-51) and (7-31).

The eigenvalues and eigenfunctions of the angular momentum operator  $L^2$  were derived in Chapters 6 and 7; the result was described in Eq. (7-30), and we may write it as

$$L^2 \Phi_J(\theta, \phi) = \hbar^2 J(J+1) \Phi_J(\theta, \phi) \quad (12-22)$$

The eigenfunctions  $f_n$  of the Schrödinger equation (12-18) may therefore be represented as

$$f_n(X, Y, Z) = g_{n,J}(R) \Phi_J(\theta, \phi) \quad (12-23)$$

Substitution into Eq. (12-18) leads to the following equation:

$$\begin{aligned}
 [H_{\text{nuc1}}(R) + \varepsilon_n(R)]f_n(X, Y, Z) &= -\frac{\hbar^2}{2\mu} \left( \frac{\partial^2}{\partial R^2} + \frac{2}{R} \frac{\partial}{\partial R} \right) g(R)\Phi_J(\theta, \varphi) \\
 &\quad + \frac{L^2}{2\mu R^2} g(R)\Phi_J(\theta, \varphi) \\
 &\quad + \left( \frac{Z_a Z_b e^2}{R} + \varepsilon_n(R) \right) g(R)\Phi_J(\theta, \varphi) \\
 &= E g(R)\Phi_J(\theta, \varphi)
 \end{aligned} \tag{12-24}$$

By making use of Eq. (12-22) this may be simplified to

$$-\frac{\hbar^2}{2\mu} \left( \frac{\partial^2}{\partial R^2} + \frac{2}{R} \frac{\partial}{\partial R} \right) g(R) + U_n(R)g(R) = E g(R) \tag{12-25}$$

with

$$U_n(R) = \varepsilon_n(R) + \frac{Z_a Z_b e^2}{R} + \frac{\hbar^2 J(J+1)}{2\mu R^2} \tag{12-26}$$

We introduce a final small adjustment by substituting

$$g(R) = \frac{\psi(R)}{R} \tag{12-27}$$

into Eq. (12-25), which leads to

$$-\frac{\hbar}{2\mu} \frac{\partial^2 \psi}{\partial R^2} + U_n(R)\psi(R) = E\psi(R) \tag{12-28}$$

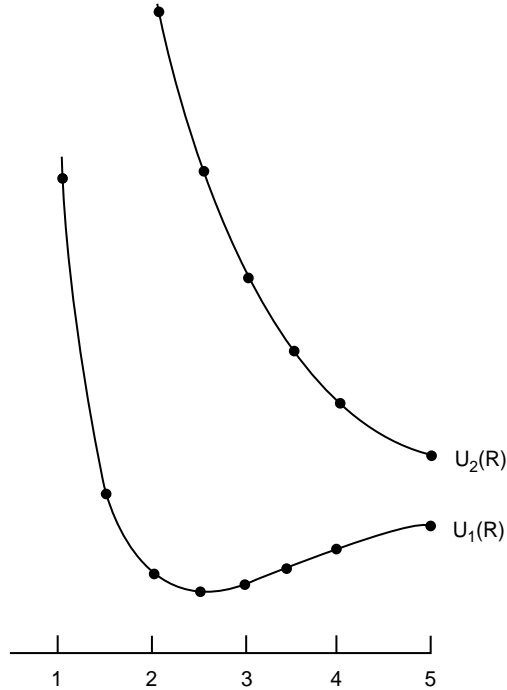
This is the customary form of the Schrödinger equation representing the vibrational motion of a diatomic molecule.

In order to solve the vibrational Schrödinger equation (12-28), it is necessary to have a general understanding of the behavior of the functions  $U_n(R)$ , which are called the molecular potential curves. We first consider the molecular ground state, and we have sketched a typical potential curve  $U_1(R)$  of a stable diatomic molecule in Figure 12-2. The potential function has a minimum

$$U_1(R_1) = \varepsilon_1(R_1) + \frac{Z_a Z_b e^2}{R_1} + \frac{\hbar^2 J(J+1)}{2\mu R_1^2} \tag{12-29}$$

at its equilibrium nuclear distance  $R_1$ . Strictly speaking, the potential function and the position of its minimum also depend on the value of its rotational quantum





**Figure 12-2** Potential curves of a diatomic molecule. The curves are derived from a calculation on the hydrogen molecular ion.

number  $J$ , but this constitutes only a very small correction and it is usually ignored. The potential function tends to infinity when  $R$  approaches zero due to the Coulomb repulsion of the two nuclei. For very large values of  $R$  the potential curve asymptotically approaches a constant value

$$U_1(\infty) = U_1(R_1) + D \tag{12-30}$$

where  $D$  is the dissociation energy of the molecule.

The solution of the Schrödinger equation (12-28) for the electronic ground state  $n = 1$  may now be derived by expanding the potential function  $U_1(R)$  as a power series in terms of its coordinate around its minimum  $R = R_1$ ,

$$U_1(R) = U_1(R_1) + \frac{1}{2}k q^2 + \frac{1}{6}k_3 q^3 + \dots \tag{12-31}$$

where

$$k = \left( \frac{\partial^2 U_1(R)}{\partial R^2} \right)_{R_1} \quad k_n = \left( \frac{\partial^n U_1(R)}{\partial R^n} \right)_{R_1} \quad q = R - R_1 \tag{12-32}$$

The customary approximation consists of terminating the power series after its second term, which leads to

$$\left(-\frac{\hbar^2}{2\mu} \frac{d^2}{dq^2} + \frac{1}{2}k q^2\right)\psi = \varepsilon\psi \quad (12-33)$$

where

$$\varepsilon = E - U_1(R_1) \quad (12-34)$$

We note that this equation (12-33) is identical to the Schrödinger equation (6-50) of the harmonic oscillator, which we discussed in Section 6.IV. Its eigenvalues  $\varepsilon_v$  are given by

$$\varepsilon_v = (v + 1/2) \hbar\omega \quad v = 1, 2, 3, \text{ etc.} \quad (12-35)$$

$$\omega = \sqrt{\frac{k}{\mu}} \quad (12-36)$$

according to Eqs. (6-62) and (6-65). The corresponding eigenfunctions  $\psi_v(q)$  are described by Eqs. (6-63) and (6-66). Finally, by combining Eqs. (12-35), (12-34), and (12-29), we find that the molecular energies  $E(1, v, J)$  are given by

$$E(1, v, j) = \varepsilon_1(R_1) + \frac{Z_a Z_b e^2}{R_1} + (v + 1/2) \hbar\omega + \frac{J(J + 1)\hbar^2}{2\mu R_1^2} \quad (12-37)$$

This is consistent with the empirical result of Eq. (12-15) if we define the rotational constant as

$$B_1 = \frac{\hbar^2}{2\mu R_1^2} \quad (12-38)$$

Since the reduced mass of a diatomic molecule is known, the equilibrium ground state internuclear distance is easily derived from experimental results of the rotational constant  $B$ . It may be instructive to present some numerical results, and we have listed the rotational constants  $B$  and ground state internuclear distances  $R_1$  for some diatomic molecules in Table 12-1. The customary unit for the rotational constant  $B$ ,  $\text{cm}^{-1}$ , refers to the wave number  $\sigma$ , which is the inverse of the wavelength of the corresponding transition:

$$E = h\nu = hc \sigma \quad (12-39)$$

It may be seen that the rotational constants  $B_1$  depend primarily on the reduced masses of the nuclei, so that their values for the hydrogen molecule and the various

**TABLE 12-1. Values of Rotational Constants  $B_1$  and Vibrational Frequencies  $\nu$  (in  $\text{cm}^{-1}$ ) and Equilibrium Internuclear Distances  $R_1$  (in  $\text{\AA}$ ) of the Ground States of Some Diatomic Molecules**

Molecule	$B_1$	$\nu$	$R_1$
H <sub>2</sub>	60.809	4395.2	0.7417
HD	45.655	3817.1	0.7414
D <sub>2</sub>	30.429	3118.5	0.7416
LiH	7.5131	1405.6	1.595
HF	20.939	4138.5	0.9171
HCl	10.591	2989.7	1.275
HBr	8.473	2649.7	1.414
HI	6.551	2309.5	1.604
Li <sub>2</sub>	0.6727	351.4	2.672
N <sub>2</sub>	2.010	2359.6	1.094
CO	1.9314	2170.2	1.128
O <sub>2</sub>	1.4457	1580.4	1.207

hydrides are significantly larger than their values for other diatomics. The values of  $R_1$  depend partially on the size of the atoms but also on the strength of the bond; this may explain the rather large values for both Cl<sub>2</sub> and Li<sub>2</sub>.

In Table 12-1 we also list the values of the vibrational frequencies  $\nu$ , again in terms of  $\text{cm}^{-1}$ . The latter depend on the values of the force constants  $k$ , which are related to the strength of the chemical bonds and, to a lesser extent, the nuclear masses. This explains the large value for HF, which has a very strong bond, and the small values for LiH and Li<sub>2</sub>, whose chemical bonds are much weaker.

Finally, we comment briefly on the potential curves corresponding to the various electronic eigenstates of the molecules. These may be divided into two different types, both sketched in Figure 12-2. The first type, denoted by  $U_1(R)$ , has a minimum at an internuclear distance  $R_1$ . The second type, denoted by  $U_2(R)$ , does not have a minimum. Excitation from the ground state to this electronic eigenstate therefore leads to molecular dissociation.

#### IV. THE HYDROGEN MOLECULAR ION

In the introduction to this chapter, we mentioned that the major goal of quantum chemistry is to determine the molecular wave function. It is hardly ever possible to derive exact molecular wave functions by solving the Schrödinger equation, but highly accurate wave functions have been obtained by making use of the variational principle discussed in Section 9.II.

The variational principle is based on the assumption that any function, in particular any molecular eigenfunction, may be expanded in terms of a known infinite complete set of functions that satisfy the same boundary conditions. It follows then that it is possible in principle to derive exact molecular eigenfunctions by

expanding the unknown eigenfunctions in terms of an infinite complete set of functions and by making use of the variational principle. In practice, calculations are restricted to finite sets that are obtained by truncating the eigenfunction expansions. The accuracy of the results is determined by two factors: the size of the set of functions used in the expansion and the convergence of the expansion. It should be obvious that for a given amount of computational effort, more accurate results can be obtained for small molecules than for large systems.

Not surprisingly, many calculations have been performed on the smallest molecules, namely, the hydrogen molecule and the hydrogen molecular ion. In those cases it is possible to obtain quite accurate results, but these two systems have also proved useful for evaluating different approaches and different wave function expansions. We first discuss the molecular ion; the hydrogen molecule will be studied in the following section.

The hydrogen molecular ion consists of two hydrogen nuclei, *a* and *b*, separated by a distance *R*, and of one electron. As our basis functions we take the hydrogen (*1s*) functions  $s_a$  and  $s_b$ , centered on nucleus *a* or *b*, respectively,

$$s_a = \frac{1}{\sqrt{\pi}} \exp(-r_a) \quad s_b = \frac{1}{\sqrt{\pi}} \exp(-r_b) \quad (12-40)$$

The molecular ion has a plane of symmetry perpendicular to the molecular axis, and according to Section 6.II, the molecular eigenfunctions should be either symmetric or antisymmetric relative to this plane. It follows therefore that the eigenfunctions derived from the basis set of two functions (12-40) are

$$\psi_1 = s_a + s_b \quad \psi_2 = s_a - s_b \quad (12-41)$$

Here  $\psi_1$  corresponds to the ground state and  $\psi_2$  to the first excited state.

The molecular eigenfunction  $\psi_1$  may also be derived from the chemical resonance principle. We recognize two possible structures of the hydrogen molecular ion. In the first structure, represented by  $s_a$ , the electron is centered on nucleus *a*, and in the second structure, represented by  $s_b$ , the electron is centered on nucleus *b*. According to the resonance principle, the molecular wave function is then obtained as a linear combination of the wave functions of the various resonance structures. In the present situation this leads to the function  $\psi_1$  of Eq. (12-41) since the two resonance structures have equal probability.

It is possible to calculate the expectation values of the Hamiltonian with respect to the wave functions  $\psi_1$  and  $\psi_2$  of (12-41). The molecular Hamiltonian is given by

$$H = -\frac{1}{2}\Delta - \frac{1}{r_a} - \frac{1}{r_b} \quad (12-42)$$

where we have used atomic units of length  $a_0$  and energy ( $e^2/a_0$ ). We have

$$H s_a = -\frac{1}{2}s_a - \frac{s_a}{r_b} \quad H s_b = -\frac{1}{2}s_b - \frac{s_b}{r_a} \quad (12-43)$$

since  $s_a$  and  $s_b$  are hydrogen atom eigenfunctions. The expectation values  $E_1$  and  $E_2$  are now

$$\begin{aligned} E_1 &= \frac{\langle \psi_1 | H | \psi_1 \rangle}{\langle \psi_1 | \psi_1 \rangle} = -\frac{1}{2} - \frac{I + J}{1 + S} \\ E_2 &= \frac{\langle \psi_2 | H | \psi_2 \rangle}{\langle \psi_2 | \psi_2 \rangle} = -\frac{1}{2} - \frac{I - J}{1 - S} \end{aligned} \quad (12-44)$$

where the three integrals  $S$ ,  $I$ , and  $J$  are defined as

$$\begin{aligned} S &= \langle s_a | s_b \rangle = \frac{1}{\pi} \int \exp(-r_a - r_b) \, d\mathbf{r} \\ I &= \langle s_a | r_b^{-1} | s_a \rangle = \frac{1}{\pi} \int r_b^{-1} \exp(-2r_a) \, d\mathbf{r} \\ J &= \langle s_a | r_a^{-1} | s_b \rangle = \frac{1}{\pi} \int r_a^{-1} \exp(-r_a - r_b) \, d\mathbf{r} \end{aligned} \quad (12-45)$$

The three integrals are calculated by introducing elliptical coordinates (see Figure 12-3)

$$\begin{aligned} \mu &= \frac{r_a + r_b}{R} & \nu &= \frac{r_a - r_b}{R} \\ 1 \leq \mu \leq \infty & & -1 \leq \nu \leq 1 \end{aligned} \quad (12-46)$$

Three dimensional integrals are transformed as follows:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) \, dx \, dy \, dz = \frac{R^3}{8} \int_0^{2\pi} d\phi \int_{-1}^1 d\nu \int_1^{\infty} f(\mu, \nu, \phi) (\mu^2 - \nu^2) \, d\mu \quad (12-47)$$

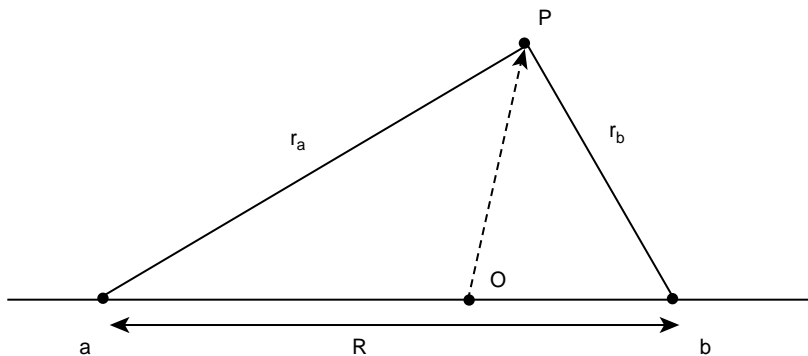


Figure 12-3 Definition of elliptical coordinates.

It is now easily found that

$$\begin{aligned}
 S &= \frac{R^3}{4} \int_{-1}^1 dv \int_1^{\infty} (\mu^2 - v^2) \exp(-\mu R) d\mu \\
 &= \left(1 + R + \frac{R^2}{3}\right) \exp(-R) \\
 I &= \frac{R^2}{2} \int_{-1}^1 dv \int_1^{\infty} (\mu + v) \exp[-R(\mu + v)] d\mu \\
 &= \frac{1}{R} [1 - (1 + R) \exp(-2R)] \\
 J &= \frac{R^2}{2} \int_{-1}^1 dv \int_1^{\infty} (\mu + v) \exp(-\mu R) d\mu \\
 &= (1 + R) \exp(-R)
 \end{aligned} \tag{12-48}$$

By substituting these results into Eq. (12-44), we obtain analytical expressions for the two expectation values  $E_1(R)$  and  $E_2(R)$  as a function of the internuclear distance  $R$ . We have sketched the two potential functions in Figure 12-2 because we used the results of this calculation to construct the potential curves  $U_1(R)$  and  $U_2(R)$  of Figure 12.2.

The result is that the lower curve  $U_1(R)$  has a minimum for  $R = 2.5$  atomic units and the energy minimum is  $-0.5648$  atomic unit. This corresponds to a molecular dissociation energy  $D_1 = 1.76$  eV and an equilibrium nuclear distance of  $1.32 \text{ \AA}$ . The potential curve  $U_2(R)$  does not have a minimum.

We list these results together with the corresponding experimental values in Table 12-2. We should realize that we have used the simplest possible wave function to obtain our theoretical values and that more accurate theoretical predictions may be derived from more elaborate variational wave functions. We discuss various possibilities.

**TABLE 12-2. Values of Dissociation Energies  $D$  (in eV) and Equilibrium Nuclear Distances  $R_1$  (in  $\text{\AA}$ ) of the Hydrogen Molecular Ion Derived from Various Variational Wave Functions**

Wave function	$D$	$R_1$
Eq. (12-41)	1.76	1.32
Eq. (12-49)	2.25	1.06
Eq. (12-50)	2.71	1.06
Eq. (12-51)	2.17	1.06
Eq. (12-52)	2.785	1.06
Experimental	2.79	1.06

We note that the electron is subject to the attractive force of two nuclei rather than just one. We may account for this by writing the wave functions as

$$\begin{aligned}\psi_1 &= s'_a + s'_b \\ s'_a &= (q^3/\pi)^{1/2} \exp(-qr_a), \dots \text{etc.}\end{aligned}\tag{12-49}$$

Here  $q$  is an effective nuclear charge, and its value is determined by making use of the variational principle. A further improvement may be obtained by writing the wave function as

$$\begin{aligned}\psi_1 &= s''_a + s''_b \\ s''_a &= (1 + \lambda z_a) \exp(-qr_a), \dots \text{etc.}\end{aligned}\tag{12-50}$$

This allows the electronic charge to shift towards the other nucleus; this effect is called polarization. We also list the theoretical results obtained from the wave functions (12-49) and (12-50) in Table 12-2. The agreement with experimental findings is much improved.

The most efficient approximation to the ground state eigenfunction makes use of elliptical coordinates. For example, the (unnormalized) wave function.

$$\psi = \exp(-q\mu)\tag{12-51}$$

produces a better result than the simple function (12-41), and the more elaborate function

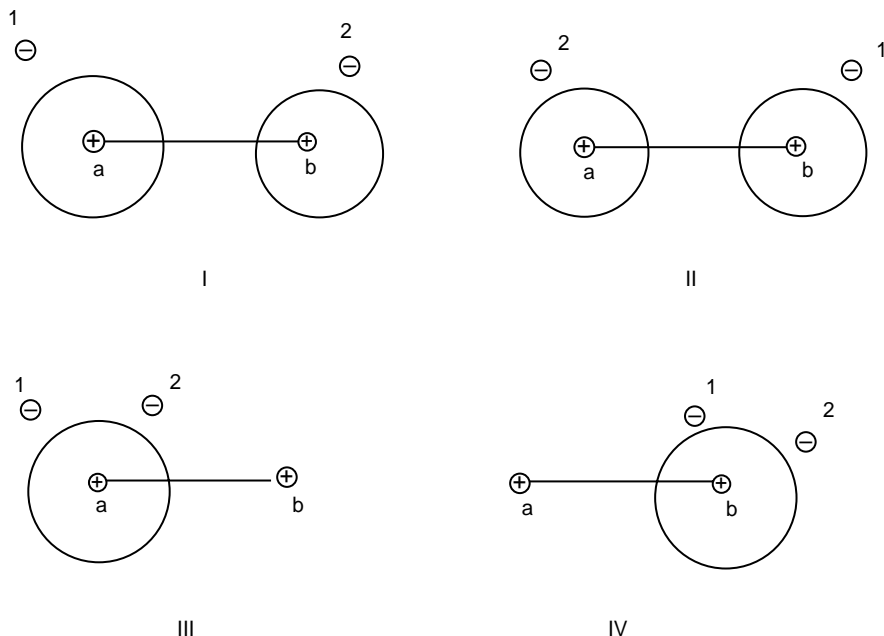
$$\psi = (1 + av^2) \exp(-q\mu)\tag{12-52}$$

leads to very good agreement with experimental findings, as we show in Table 12-2.

It is possible to derive highly accurate theoretical results for the hydrogen molecular ion. In fact, its Schrödinger equation has been solved exactly. However, the analytical form of this solution is quite complex and difficult to use, so it has found few practical applications. Similarly, the use of elliptical coordinates may be convenient for the hydrogen molecular ion or even the hydrogen molecule, but it is much less convenient for more complex molecules and its practical usefulness in molecular calculations is therefore limited.

## V. THE HYDROGEN MOLECULE

A chemical bond between two atoms usually involves a pair of electrons. Since the hydrogen molecule is the smallest and simplest molecule, it is a useful system to illustrate the quantum mechanical description of the chemical bond. Initially this description was based on two different models that became known as the valence-bond (VB) model and the molecular-orbital (MO) model. We explain



**Figure 12-4** Various resonance structures of the hydrogen molecule.

both models for the hydrogen molecule by expressing its wave function in terms of the simple hydrogen  $s_a$  and  $s_b$  orbitals that we defined in Eq. (12-40).

Just as in the case of the hydrogen molecular ion, we may approximate the molecular wave function by writing it as a linear combination of the various resonance structures sketched in Figure 12-4. It is easily seen that structure I represents the situation where electron 1 is centered on nucleus a and electron 2 is centered on nucleus b. The corresponding wave function may be represented as

$$\psi_{\text{I}}(1, 2) = s_a(1)s_b(2) \quad (12-53)$$

There is, of course, an equal probability of finding electron 1 on nucleus b and electron 2 on nucleus a. This corresponds to resonance structure II, which is represented by the function

$$\psi_{\text{II}}(1, 2) = s_b(1)s_a(2) \quad (12-54)$$

The complete molecular eigenfunction, including the electron spins, is now given by

$$\Psi_{\text{VB}} = [s_a(1)s_b(2) + s_b(1)s_a(2)][\alpha(1)\beta(2) - \beta(1)\alpha(2)] \quad (12-55)$$



As the name indicates, this is the VB molecular wave function. Since it must be antisymmetric with respect to permutations, it contains a singlet spin function.

In the MO model we write the molecular wave function as a product of two one-electron functions or orbitals

$$\Psi_{\text{MO}} = \psi_1(1)\psi_1(2)[\alpha(1)\beta(2) - \beta(1)\alpha(2)] \quad (12-56)$$

consistent with the Hartree-Fock approximation. We may then approximate the molecular orbital  $\psi_1$  further as

$$\psi_1(1) = s_a(1) + s_b(1) \quad (12-57)$$

by analogy with Eq. (12-41) of the hydrogen molecular ion.

The difference between the VB and MO functions may best be illustrated by substituting Eq. (12-57) into Eq. (12-56) and by writing out the product

$$\Psi_{\text{MO}} = s_a(1)s_a(2) + s_a(1)s_b(2) + s_b(1)s_a(2) + s_b(1)s_b(2) \quad (12-58)$$

where we have left out the spin functions. It is easily seen that this function represents a superposition of all four of the resonance structures sketched in Figure 12-4, namely, the two ionic structures, III and IV, with both electrons centered on either nucleus a or nucleus b, in addition to the two nonionic structures, I and II, that we considered earlier.

In constructing a molecular wave function, we should in principle consider all possible resonance structures, and our VB function fails to include the ionic resonance structures III and IV. On the other hand, the MO function assigns the same probability to the ionic structures III and IV as to the nonionic structures I and II. In short, the VB function ignores the effect of the ionic structures, while the MO function overestimates their contribution.

It seems logical to determine the true contribution of the ionic structures by means of the variational principle. This may be accomplished by proposing the molecular wave function

$$\Psi = \rho[s_a(1)s_b(2) + s_b(1)s_a(2)] + \sigma[s_a(1)s_a(2) + s_b(1)s_b(2)] \quad (12-59)$$

and by subsequent minimization with respect to the parameters  $\rho$  and  $\sigma$ .

The theoretical results for the molecular dissociation energies  $D$  and internuclear distances  $R_1$  are significantly improved if we substitute the atomic  $s'_a$  and  $s'_b$  functions of Eq. (12-49) that contain an effective nuclear charge rather than the exact hydrogen atom functions. We list the results of the various variational calculations in Table 12-3. We have also listed the results of an exact Hartree-Fock calculation and the experimental values. It is interesting to note that the Hartree-Fock method leads to less accurate predictions than the VB method. This is due to the fact that the VB wave function takes electron repulsion into account by always assigning the electrons to different atoms, while the Hartree-Fock function does not. We also

**TABLE 12-3. Values of the Hydrogen Molecule Dissociation Energies  $D$  (in eV) and Equilibrium Nuclear Distances  $R_1$  (in Å) Derived from Various Wave Functions**

Wave Function	$D$	$R_1$
Eq. (12-55) VB, $s$	3.14	0.869
Eq. (12-55) VB, $s'$	3.78	0.743
Eq. (12-56) MO, $s$	2.70	0.85
Eq. (12-56) MO, $s'$	3.49	0.732
Eq. (12-59) $s'$	4.02	0.749
HF	3.62	0.74
“Exact”	4.7467	0.741
Experimental	4.747	0.741

report the results of a very elaborate variational calculation (which we denote by “Exact”) without describing the details of the calculation. These results agree quite closely with the corresponding experimental values.

The VB model is more compatible with the chemical understanding of valence than the MO model, and it also leads to more accurate theoretical predictions. Unfortunately, it is not as well suited as the MO model in combination with the Hartree-Fock method for quantum mechanical calculations on larger molecules. Consequently, the majority of molecular structure calculations are now based on the MO rather than the VB model.

## VI. THE CHEMICAL BOND

In the introduction to this chapter, we mentioned that the determination and interpretation of molecular structure is one of the major goals of quantum chemistry. In many respects, this goal has been achieved because there are now sophisticated computer programs available that are capable of producing accurate molecular wave functions for fairly complex molecules. However, a detailed description of these programs falls outside the scope of this book. Instead we outline the general behavior of the wave functions that are associated with a chemical bond.

It is generally accepted in chemistry that a single bond between two atoms, A and B, is formed by a pair of electrons with opposite spins. We may describe this situation in the MO representation by a wave function

$$\Psi_{\text{MO}} = [t_{\text{A}}(1) + \rho t_{\text{B}}(1)][t_{\text{A}}(2) + \rho t_{\text{B}}(2)] \quad (12-60)$$

Here  $t_{\text{A}}$  and  $t_{\text{B}}$  are so-called atomic valence orbitals centered on nucleus A and nucleus B, respectively. For convenience, we have omitted the singlet spin function.

Different atoms have different electron-attracting powers, and this effect is accounted for by the introduction of the parameter  $\rho$  in Eq. (12-60). The parameter  $\rho$  is smaller than unity if atom A has the greater attractive power for electrons, and it is larger than unity in the opposite case.

We may also represent a chemical bond by means of a VB wave function

$$\begin{aligned}\Psi_{\text{VB}} = & t_{\text{A}}(1)t_{\text{B}}(2) + t_{\text{B}}(1)t_{\text{A}}(2) \\ & + \lambda t_{\text{A}}(1)t_{\text{A}}(2) + \mu t_{\text{B}}(1)t_{\text{B}}(2)\end{aligned}\quad (12-61)$$

This function is a superposition of a nonionic structure and of two ionic structures. The values of the two parameters  $\lambda$  and  $\mu$  depend again on the relative electron-attracting powers of the two atoms A and B.

Approximate theoretical expressions for the atomic valence orbitals  $t_{\text{A}}$  and  $t_{\text{B}}$  may be derived by expanding in terms of atomic orbitals and by determining the expansion coefficients from the variational principle. Theoretical considerations show that there is a correlation between the strength of a chemical bond and the values of the overlap integral  $S_{\text{AB}}$  between the two atomic valence orbitals  $t_{\text{A}}$  and  $t_{\text{B}}$ :

$$S_{\text{AB}} = \langle t_{\text{A}} | t_{\text{B}} \rangle \quad (12-62)$$

The strength and energy of the chemical bond between the atoms A and B are directly related to the value of  $S_{\text{AB}}$ ; the higher the value of  $S_{\text{AB}}$ , the stronger the bond. It therefore becomes possible to derive the analytical form of  $t_{\text{A}}$  and  $t_{\text{B}}$  just by maximizing the overlap integral rather than by minimizing the energy according to the variational principle.

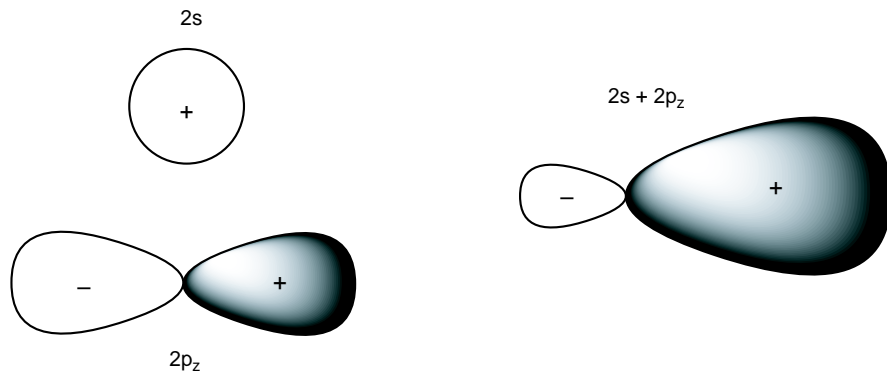
In order to illustrate the above ideas, we derive an approximate expression for the wave function of the diatomic  $\text{N}_2$  molecule. Table 11-1 shows that the nitrogen atom has seven electrons. Two electrons are assigned to the ( $1s$ ) orbital, which we denote by  $h$  and which does not participate in the chemical bond. The remaining five electrons are distributed over the ( $2s$ ) and ( $2p$ ) orbitals, which we denote by  $s$ ,  $p_x$ ,  $p_y$ , and  $p_z$  and which are the basis for the chemical bonds in the  $\text{N}_2$  molecule.

Figure 12-5 shows how the strongest chemical bond may be constructed from the atomic  $s$  and  $p_z$  orbitals if we take the  $Z$  axis along the molecular axis. It may be seen that the function

$$t_{\text{A}} = s_{\text{A}} + p_{z\text{A}} \quad (12-63)$$

is concentrated around the  $Z$  axis and points in the positive  $z$  direction, while the orbital

$$t'_{\text{A}} = s_{\text{A}} - p_{z\text{A}} \quad (12-64)$$



**Figure 12-5** A directed valence orbital of the nitrogen atom that is obtained by combining a  $2s$  and a  $2p$  orbital.

points in the negative  $z$  direction. These orbitals are known as directed valence orbitals or as hybridized orbitals. The overlap integral

$$\langle t_A | t'_B \rangle \quad (12-65)$$

is around 0.7 and the bonding orbital

$$\sigma = t_A + t'_B \quad (12-66)$$

represents a strong nitrogen-nitrogen bond.

Figure 12-5 shows that the orbitals  $p_{xA}$  and  $p_{xB}$  and the orbitals  $p_{yA}$  and  $p_{yB}$  may also be combined to form bonding orbitals:

$$\pi = p_{xA} + p_{xB} \quad \pi' = p_{yA} + p_{yB} \quad (12-67)$$

but in this case the overlap integral is much smaller than for the  $\sigma$  bonding orbital and the resulting nitrogen-nitrogen  $\pi$  bonds are therefore weaker.

The orbital  $t'_A$  has a well-defined direction, and it is known as a lone pair orbital. It does not participate in any chemical bond since it points in a direction away from the other nucleus, B. It is basically a nonbonding atomic orbital with a corresponding energy.

We should also mention the existence of antibonding orbitals:

$$\begin{aligned} \bar{\sigma} &= t_A - t'_B \\ \bar{\pi} &= p_{xA} - p_{xB} \quad \bar{\pi}' = p_{yA} - p_{yB} \end{aligned} \quad (12-68)$$

While the bonding orbitals have positive overlap charges and energies that are lower than their corresponding atomic energies, the antibonding orbitals have

negative overlap integrals and energies that are actually higher than the corresponding atomic energies.

It may be seen that the bonding orbitals have the lowest energies, the antibonding orbitals have the highest energies, and the nonbonding or lone pair orbitals have energies in between. We may therefore rank the energies of the various orbitals as follows:

$$\varepsilon(1s_A), \varepsilon(1s_B) < \varepsilon(\sigma) < \varepsilon(\pi), \varepsilon(\pi') < \varepsilon(t'_A), \varepsilon(t'_B) < \varepsilon(\bar{\pi}) < (\bar{\sigma}) \quad (12-69)$$

The electronic configuration of the nitrogen molecule ground state is therefore

$$(1s_A)^2(1s_B)^2(\sigma)^2(\pi)^2(\pi')^2(t'_A)^2(t'_B)^2 \quad (12-70)$$

since the molecule has 14 electrons. The molecule has a triple bond consisting of one  $\sigma$  bond and two  $\pi$  bonds, and it also has two electron pairs in the two lone pair orbitals. We leave it to the reader to construct the corresponding antisymmetric wave function including the electron spins. Admittedly, this function constitutes no more than an educated guess as to the molecular wave function, but it offers a good understanding of the molecular electronic structure. It has also been used as a basis for various quantum mechanical calculation on the nitrogen molecule that have produced surprisingly good results.

## VII. THE STRUCTURES OF SOME SIMPLE POLYATOMIC MOLECULES

The above simple model of the chemical bond may also be applied to polyatomic molecules. We consider a few molecules as examples, beginning with the methane molecule,  $\text{CH}_4$ . This molecule has four equivalent carbon-hydrogen bonds that form a tetrahedral geometry. We again use the molecular orbital model, and we assume that the bond orbitals may be constructed from atomic valence orbitals centered on the carbon atom and from the four hydrogen ( $1s$ ) orbitals. The ground state electronic configuration of the carbon atom is  $(1s)^2(2s)^2(2p)^2$ , but we excite this configuration to  $(1s)^2(2s)(2p)^3$  in order to obtain the valence orbitals since the excitation energy is more than compensated for by the improvement in bond energies.

The atomic valence orbitals may now be derived from the following guidelines:

1. They must be linear combinations of the  $2s$ ,  $2p_x$ ,  $2p_y$ , and  $2p_z$  orbitals.
2. They must be orthogonal.
3. They must be equivalent.
4. Their overall electronic distribution must correspond to the configuration  $(2s)(2p)^3$ .

It is easily verified that the following set of valence orbitals meets all four of the above requirements:

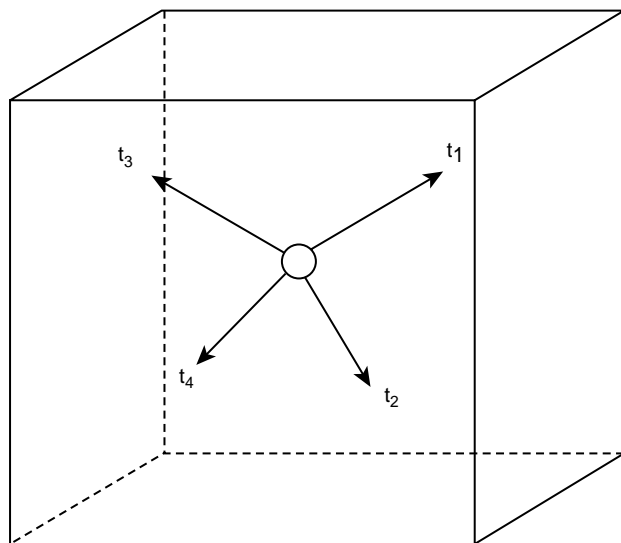
$$\begin{aligned}
 t_1 &= \frac{1}{2}(s + p_x + p_y + p_z) \\
 t_2 &= \frac{1}{2}(s + p_x - p_y - p_z) \\
 t_3 &= \frac{1}{2}(s - p_x + p_y - p_z) \\
 t_4 &= \frac{1}{2}(s - p_x - p_y + p_z)
 \end{aligned}
 \tag{12-71}$$

where we have denoted the ( $2s$ ) orbital by the symbol  $s$ .

We show the directions of these four atomic valence orbitals in Figure 12-6, and it may be seen that they exhibit a tetrahedral geometry pattern. In order to describe the wave function of the methane molecule, we place the four hydrogen atoms along the tetrahedral directions at the appropriate distances and we define the four bond orbitals

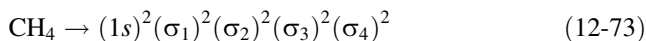
$$\sigma_i = t_i + h_i \tag{12-72}$$

where  $h_i$  are the hydrogen atom ( $1s$ ) orbitals. We assume that the carbon and hydrogen atoms have comparable electron-attracting powers so that we may set the

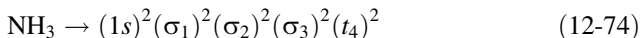


**Figure 12-6** Four  $sp^3$  hybridized valence orbitals forming a tetrahedral geometry.

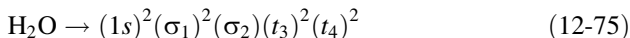
parameter  $\rho$  of Eq. (12-60) equal to unity. It follows that the electronic structure of the methane molecule may be represented by the following distribution:



It may be interesting to compare the electronic structure of the methane molecule with the structures of the ammonia molecule  $\text{NH}_3$  and the water molecule  $\text{H}_2\text{O}$  since they all have the same number of electrons. The structure of  $\text{NH}_3$  may be derived from Eq. (12-73) by removing one hydrogen nucleus and by replacing the bond orbital  $\sigma_4$  by the corresponding atomic valence orbital  $t_4$ :



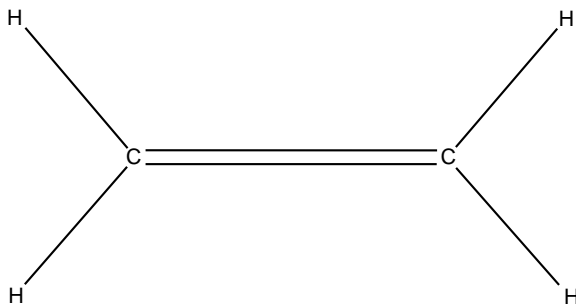
The electronic structure of the water molecule  $\text{H}_2\text{O}$  may be obtained in a similar fashion by replacing two bond orbitals,  $\sigma_3$  and  $\sigma_4$ , by atomic valence orbitals:



It should be noted that the charge clouds associated with the lone pair electrons in ammonia and water point in well-defined directions that form an approximate tetrahedral geometry pattern. The lone pair electrons may therefore give rise to electric interactions with other molecules that are important in biochemistry and medicine.

The four atomic valence orbitals of Eq. (12-71) are obtained as hybrids of one atomic ( $2s$ ) orbital and three atomic ( $2p$ ) orbitals, namely, ( $2p_x$ ), ( $2p_y$ ), and ( $2p_z$ ), and they are known as a set of  $sp^3$  hybridized orbitals. The  $sp^3$  type is the most common hybridization type, but there are two alternative schemes for constructing atomic valence orbitals, namely,  $sp^2$  and  $sp$  hybridization.

The  $sp^2$  hybridization pattern is best explained by considering the structure of the ethylene molecule,  $\text{C}_2\text{H}_4$  (see Figure 12-7). This molecule has a planar structure and the four carbon-hydrogen bonds form  $120^\circ$  angles with the carbon-carbon bond, which is assumed to be a double bond. Since all the bonds are located in



**Figure 12-7** Geometry of the ethylene molecule.

the molecular plane, which we take as the  $XY$  plane, they must be represented as linear combinations of the carbon ( $2s$ ) orbital and the carbon ( $2p_x$ ) and ( $2p_y$ ) orbitals, with exclusion of the carbon ( $2p_z$ ) orbital.

The three atomic valence orbitals must again be equivalent and orthogonal. If we denote the carbon atoms by A and B, then the three  $sp^2$  hybridized valence orbitals of carbon atom A are given by

$$\begin{aligned} t_{A1} &= (1/\sqrt{3})s_A + (\sqrt{2}/\sqrt{3})p_{xA} \\ t_{A2} &= (1/\sqrt{3})s_A - (1/\sqrt{6})p_{xA} + (1/\sqrt{2})p_{yA} \\ t_{A3} &= (1/\sqrt{3})s_A - (1/\sqrt{6})p_{xA} - (1/\sqrt{2})p_{yA} \end{aligned} \quad (12-76)$$

The valence orbitals of atom B are defined in a similar fashion.

Two of the atomic valence orbitals on carbon atoms A and B may now be combined to form a carbon-carbon  $\sigma$  bond, and the other carbon valence orbitals are combined with the four hydrogen ( $1s$ ) orbitals to form carbon-hydrogen bonds. The two carbon  $p_z$  orbitals  $p_{zA}$  and  $p_{zB}$  may be combined to form a carbon  $\pi$  bond, as illustrated in Figure 12-8. The  $sp^2$  hybridization in the ethylene molecule therefore predicts a carbon-carbon double bond consisting of a  $\sigma$  bond and an additional  $\pi$  bond.

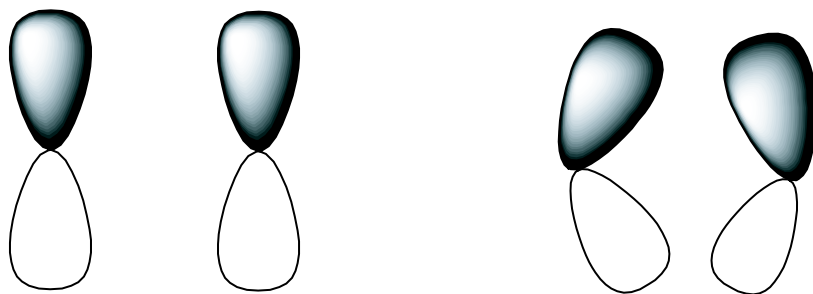
The simplest molecule exhibiting  $sp$  hybridization is acetylene,  $C_2H_2$ , which has a linear structure and a triple carbon-carbon bond. The triple bond consists of one  $\sigma$  bond and two additional  $\pi$  bonds.

If we take the  $X$  axis as the molecular axis and denote the carbon atoms again by A and B, then the four atomic valence orbitals are obtained as linear combinations of the ( $2s$ ) and ( $2p_x$ ) orbitals:

$$\begin{aligned} t_{A1} &= s_A + p_{xA} & t_{B1} &= s_B - p_{xB} \\ t_{A2} &= s_A - p_{xA} & t_{B2} &= s_B + p_{xB} \end{aligned} \quad (12-77)$$

The orbitals  $t_{A1}$  and  $t_{B1}$  form a carbon-carbon  $\sigma$  orbital

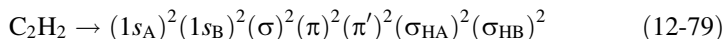
$$\sigma = t_{A1} + t_{B1} \quad (12-78)$$



**Figure 12-8** Formation of a  $\pi$  bond.



and the other  $sp$  orbitals  $t_{A2}$  and  $t_{B2}$  may be combined with the hydrogen ( $1s$ ) orbitals to form two carbon-hydrogen bonding orbitals  $\sigma_{HA}$  and  $\sigma_{HB}$ . The two orbitals  $p_{yA}$  and  $p_{yB}$  are combined to form a carbon-carbon bonding orbital  $p$ , and the two orbitals  $p_{zA}$  and  $p_{zB}$  form a similar bonding orbital  $\pi'$ . The electron structure is then given by



It is interesting to compare the electronic structure (12-79) of acetylene with the electronic structure (12-70) of the  $N_2$  molecule. The two molecules have the same number of electrons, and if we remove the two hydrogen nuclei from  $C_2H_2$  its electronic configuration becomes identical with the  $N_2$  configuration.

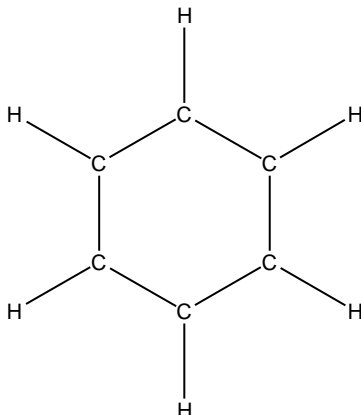
It may seem that the various wave functions discussed in this section are of a rather crude nature, but they led to surprisingly good results when used as a basis for the calculation of molecular properties. More important, they have contributed to a general understanding of molecular structure.

## VIII. THE HÜCKEL MOLECULAR ORBITAL METHOD

As a graduate student, I was once asked as part of an oral examination to explain the semiempirical Hückel method. The examiner, a professor of theoretical physics, was not convinced by my attempts to justify the various approximations that were required for its derivation. I finally said: "I may not be able to justify the Hückel method, but it works extremely well and it has been successful in interpreting and predicting many complex and sophisticated phenomena in organic chemistry." In retrospect, this does not seem a bad description of Hückel theory. Incidentally, I passed the examination.

During the 1930s, the German physicist Erich Hückel proposed a semiempirical theory to describe the electronic structure of aromatic and conjugated organic molecules. The latter are reactive compounds that had become important in a number of practical applications. For example, some aromatic compounds were starting products in the early industrial production of dyestuffs. Chemists proposed that some of the characteristic properties of aromatic and conjugated molecules could be attributed to the presence of delocalized electron orbitals. Such orbitals were not confined to a single chemical bond but could extend over a number of bonds or even the entire molecule.

Hückel's theory offered a theoretical description of these delocalized orbitals based on a number of rather drastic approximations. It is best illustrated for the benzene molecule  $C_6H_6$ , which is both the smallest aromatic and the prototype of the aromatics. The benzene molecule is planar, all 12 atoms are located in the  $XY$  plane, and its geometry is that of a regular hexagon (see Figure 12-9). It is important to note that all carbon-carbon bonds are equivalent; they all have the same length and energy. The six carbon-hydrogen bonds are also equivalent.



**Figure 12-9** Geometry of the benzene molecule.

The  $\sigma$  bonding orbitals in the plane of the molecules may all be expressed in terms of  $sp^2$  hybridized orbitals that are linear combinations of the carbon ( $2s$ ), ( $2p_x$ ), and ( $2p_y$ ) orbitals and of the hydrogen ( $1s$ ) orbitals. This bonding skeleton involves three of the four valence electrons of each carbon atom, which leaves six electrons unaccounted for. At the same time, there are six carbon ( $2p_z$ ) or  $\pi$  orbitals available. It therefore seems logical to assume that these six electrons may be distributed over the six available  $\pi$  orbitals and that their molecular orbitals  $\phi$  may be represented as linear combination of the atomic orbitals  $\pi_i$ :

$$\phi = \sum_i a_i \pi_i \quad (12-80)$$

The specific form of the expansion coefficients  $a_i$  may be derived by means of the variational principle described in Sections 9.II and 9.III and by means of Eq. (9-18) in particular. In our present situation, we have a finite basis set so that we may write the variational equations as

$$\sum_{k=1}^N (H_{jk} - \epsilon S_{jk}) a_k = 0 \quad j = 1, 2, \dots, N \quad (12-81)$$

The matrix elements  $H_{jk}$  and  $S_{jk}$  are defined as

$$H_{jk} = \langle \pi_j | H | \pi_r \rangle \quad S_{jk} = \langle \pi_j | \pi_k \rangle \quad (12-82)$$

Here  $H$  represents the effective Hamiltonian acting on an electron in a delocalized  $\pi$  orbital. It is the sum of the kinetic energy and the electrostatic interactions between a particular  $\pi$  electron and the nuclei, the  $\sigma$  electrons, and the other  $\pi$  electrons.

It will appear that its particular form does not really matter because of the semi-empirical nature of the theory.

Hückel proposed that the matrix elements  $H_{ij}$  and  $S_{ij}$  may be approximated as semiempirical parameters as follows:

$$\begin{aligned}
 H_{jk} &= \alpha & \text{if } j = k \\
 H_{jk} &= \beta & \text{if } j \text{ and } k \text{ are separated by one bond} \\
 H_{jk} &= 0 & \text{if } j \text{ and } k \text{ are separated by more than one bond} \\
 S_{jk} &= 1 & \text{if } j = k \\
 S_{jk} &= 0 & \text{if } j \neq k
 \end{aligned}
 \tag{12-83}$$

The Hückel equations may then be obtained by substituting the approximate equations (12-83) into the variational equation (12-81). In general, these are  $N$  homogeneous linear equations with  $N$  unknowns. These equations were discussed in Section 2.VIII, where we showed that they have nonzero solutions only if the determinant of the coefficients is zero. The standard procedure for solving the Hückel equations consists of first evaluating the values of the parameter  $\epsilon$  of Eq. (12-81) in order to determine the energy eigenvalues and then solving the linear equations in order to determine the corresponding eigenvectors and molecular orbitals.

In a few specific situations, the Hückel equations may be solved by a much simpler procedure in which there is no need to evaluate the determinant. The first case is a ring system containing  $N$  carbon atoms, and the second case is a conjugated hydrocarbon chain of  $N$  atoms with alternate single and double bonds.

The Hückel equations of the ring system and of the chain are similar, but there are minor differences. The equations for the ring are

$$\begin{aligned}
 \beta a_N + (\alpha - \epsilon)a_1 + \beta a_2 &= 0 \\
 \beta a_{k-1} + (\alpha - \epsilon)a_k + \beta a_{k+1} &= 0 & k = 2, 3, \dots, N-1 \\
 \beta a_{N-1} + (\alpha - \epsilon)a_N + \beta a_1 &= 0
 \end{aligned}
 \tag{12-84}$$

and those for the chain are

$$\begin{aligned}
 (\alpha - \epsilon)a_1 + \beta a_2 &= 0 \\
 \beta a_{k-1} + (\alpha - \epsilon)a_k + \beta a_{k+1} &= 0 & k = 2, 3, \dots, N-1 \\
 \beta a_{N-1} + (\alpha - \epsilon)a_N &= 0
 \end{aligned}
 \tag{12-85}$$

The set of equations (12-84) are solved by substituting

$$a_k = e^{ik\lambda} \tag{12-86}$$

The middle equations are then

$$\beta e^{-i\lambda} + (\alpha - \epsilon) + \beta e^{i\lambda} = 0 \tag{12-87}$$

and the expression (12-87) is a solution if

$$\varepsilon = \alpha + \beta(e^{i\lambda} + e^{-i\lambda}) = \alpha + 2\beta \cos \lambda \quad (12-88)$$

Substitution of Eqs. (12-86) and (12-88) into the first and last equation (12-84) gives

$$e^{iN\lambda} - 1 = 0 \quad (12-89)$$

and these equations are solved for

$$\lambda = \frac{2\pi n}{N} \quad n = 0, \pm 1, \pm 2, \quad \text{etc.} \quad (12-90)$$

The eigenvalues of Eq. (12-84) are therefore given as

$$\varepsilon_n = \alpha + 2\beta \cos(2\pi n/N) \quad n = 0, \pm 1, \pm 2, \quad \text{etc.} \quad (12-91)$$

In order to solve the Hückel equations (12-85) for the chain system we must substitute

$$a_k = Ae^{ik\lambda} + Be^{-ik\lambda} \quad (12-92)$$

This provides a solution of all equations except the first and the last one if we take

$$\varepsilon_n = \alpha + 2\beta \cos \lambda \quad (12-93)$$

We again substitute the solutions (12-92) and (12-93) into the first and last Eq. (12-85) and we obtain

$$\begin{aligned} A + B &= 0 \\ Ae^{(N+1)i\lambda} + Be^{-(N+1)i\lambda} &= 0 \end{aligned} \quad (12-94)$$

or

$$\begin{aligned} B &= -A \\ \sin[(N+1)\lambda] &= 0 \end{aligned} \quad (12-95)$$

The eigenvalues are now given by

$$\lambda = \frac{n\pi}{N+1} \quad n = 1, 2, 3, \dots, N \quad (12-96)$$

or

$$\varepsilon_n = \alpha + 2\beta \cos \frac{n\pi}{N+1} \quad (12-97)$$

The corresponding eigenvectors are

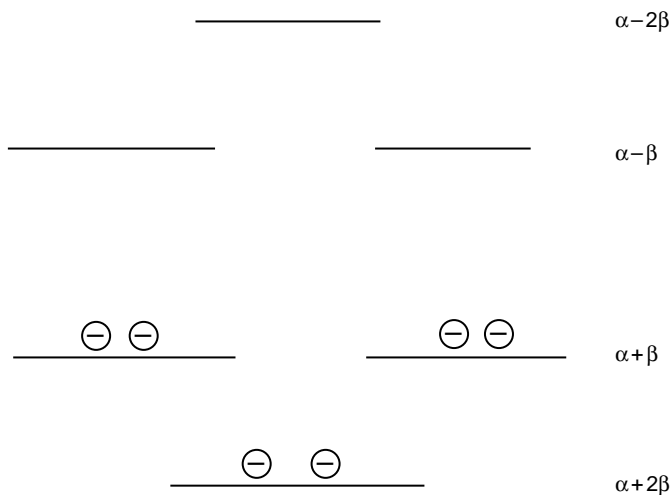
$$a_k = \sin \frac{nk\pi}{N+1} \quad (12-98)$$

It may be instructive to present a few specific examples of the above results. The benzene molecule is an aromatic ring systems of six carbon atoms, and its eigenvalues are described by Eq. (12-90) by substituting  $N = 6$ . The results are

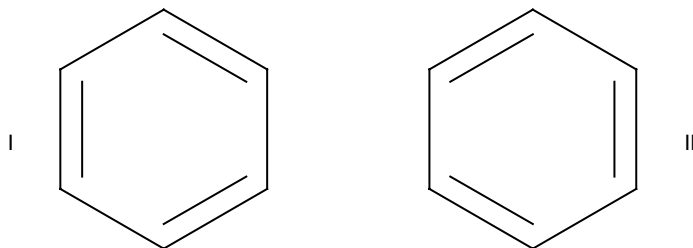
$$\begin{aligned} \varepsilon_0 &= \alpha + 2\beta \cos 0^\circ = \alpha + 2\beta \\ \varepsilon_1 &= \varepsilon_{-1} = \alpha + 2\beta \cos 60^\circ = \alpha + \beta \\ \varepsilon_2 &= \varepsilon_{-2} = \alpha + 2\beta \cos 120^\circ = \alpha - \beta \\ \varepsilon_3 &= \alpha + 2\beta \cos 180^\circ = \alpha - 2\beta \end{aligned} \quad (12-99)$$

We have sketched the energy level diagram in Figure 12-10. It should be realized that  $\beta$  is negative and that  $\varepsilon_0$  is the lowest energy eigenvalue. The next eigenvalue,  $\varepsilon_1$ , is twofold degenerate. The molecular ground state has a pair of electrons in the eigenstates  $\varepsilon_0$ ,  $\varepsilon_1$ , and  $\varepsilon_{-1}$ , and its energy is therefore

$$E(\text{benzene}) = 2(\alpha + 2\beta) + 4(\alpha + \beta) = 6\alpha + 8\beta \quad (12-100)$$



**Figure 12-10** Energy level diagram of the benzene molecule.



**Figure 12-11** Resonance structures of the benzene molecule.

It is easily shown that a localized  $\pi$  orbital has an energy  $\alpha + \beta$ , so that one of the two benzene structures of Figure 12-9 with fixed  $\pi$  bonds has an energy

$$E_1 = 6\alpha + 6\beta \quad (12-101)$$

It follows that the energy of the delocalized bond model is lower by an amount  $2\beta$  than the energy of the structure with localized  $\pi$  bonds. This energy difference is called the resonance energy of benzene.

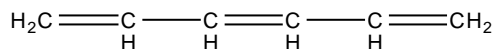
In early theoretical work on the benzene structure it was assumed that the molecule “resonated” between the two structures I and II of Figure 12-11 and that this resonance effect resulted in lowering the energy by an amount that was defined as the resonance energy. However, the molecular orbital model that we have used is better suited for numerical predictions of the resonance energies of aromatic molecules than the corresponding VB model.

As a second example, we calculate the energy eigenvalues of the hexatriene molecule. We have sketched its structure containing localized bonds in Figure 12-12; it has six carbon atoms, and the energy of the three localized  $\pi$  bonds is again  $6\alpha + 6\beta$ . The three lowest energy eigenvalues, according to the Hückel theory, may be derived from Eq. (12-97); they are

$$\begin{aligned} \varepsilon_1 &= \alpha + 2\beta \cos 25.71^\circ = \alpha + 1.8019 \beta \\ \varepsilon_2 &= \alpha + 2\beta \cos 51.71^\circ = \alpha + 1.2470 \beta \\ \varepsilon_3 &= \alpha + 2\beta \cos 77.14^\circ = \alpha + 0.4450 \beta \end{aligned} \quad (12-102)$$

The total molecular energy, according to the Hückel theory, is therefore  $6\alpha + 6.9879\beta$  and the resonance energy is  $0.9879\beta$ .

It turned out that the Hückel MO theory could be successfully applied to the prediction of molecular geometries, electronic charge densities, chemical reactivities,



**Figure 12-12** Structure of the 1,3,5 hexatriene molecule.

and a variety of other molecular properties. It therefore became quite popular, and the number of research publications in quantum chemistry based on the Hückel theory was surprisingly large. The Hückel theory was later extended to localized orbitals so that it was applicable to molecules other than aromatics or conjugated systems. The Hückel method may therefore be considered a precursor of the more sophisticated contemporary theories of molecular structure. We will not describe these latter theories; they fall outside the scope of this book.

## IX. PROBLEMS

- 12-1** Calculate the reduced nuclear masses for the molecules  $H_2$ , HD, HF and CO.
- 12-2** The rotational constant of the ground state of the HF molecule is  $B_1 = 20.939 \text{ cm}^{-1}$ . Calculate the corresponding equilibrium nuclear distance  $R_1$ .
- 12-3** In the ground state of the CO molecule the rotational constant  $B_1 = 1.9314 \text{ cm}^{-1}$  and in the first excited state the value is  $B_2 = 1.6116 \text{ cm}^{-1}$ . Calculate the equilibrium nuclear distances  $R_1$  and  $R_2$  in both electronic eigenstates.
- 12-4** The vibrational frequency  $\nu$  of the hydrogen molecule ground state is  $4395.2 \text{ cm}^{-1}$ . Assuming that the vibrational motion is harmonic, calculate the force constant  $k$  of the harmonic motion. From this result derive the expectation value of  $q^2$  where  $q = R - R_1$  represents the change in internuclear distance due to the vibrational motion. Compare the square root of  $\langle q^2 \rangle$  with the equilibrium internuclear distance  $R_1$  that is derived from the rotational constant  $B_1 = 60.809 \text{ cm}^{-1}$ .
- 12-5** Perform the same calculation for the oxygen molecules  $O_2$  where  $\nu = 1580.4 \text{ cm}^{-1}$  and  $B_1 = 1.4457 \text{ cm}^{-1}$ .
- 12-6** The rotational constants  $B_1$  of  $H_2$ , HD and  $D_2$  are  $60.809 \text{ cm}^{-1}$ ,  $45.655 \text{ cm}^{-1}$  and  $30.429 \text{ cm}^{-1}$  respectively. Calculate the differences in the equilibrium nuclear distances  $R_1$  of the three molecules
- 12-7** Derive an analytical expression for the overlap integral

$$S = \langle s_a | s_b \rangle$$

defined by Eqs. (12-40) and (12-45) and calculate its value for the internuclear distances  $R_{ab} = 2.0 \text{ a.u.}$ ,  $R_{aa} = 2.5 \text{ a.u.}$  and  $R_{ab} = 3.0 \text{ a.u.}$

- 12-8** Determine the numerical values of the normalized wave function  $(s_a + s_b)$  of the hydrogen molecular ion at the positions of the nuclei  $a$  and  $b$  and at the point midway between the two nuclei. Which of the two values is larger?

- 12-9** Explain why the VB wave function gives a lower energy for the hydrogen molecule than the MO wave function.
- 12-10** Explain why the H-N-H bond angle of  $108^\circ$  in ammonia is smaller than the  $109.5^\circ$  H-C-H bond angle of the water molecule.
- 12-11** The oxygen molecule  $O_2$  is one of the few molecules whose ground state is a triplet spin state. Explain this on the basis of the relations we presented in Eqs. (12-69) and (12-70)
- 12-12** Explain the electronic structure of the HCN molecule in terms of  $\sigma$  and  $\pi$  orbitals.
- 12-13** Solve the Hückel equations of a conjugated hydrocarbon ring system  $C_8H_8$  containing eight carbon atoms. Derive the Hückel eigenvalues and eigenfunctions and also the resonance energy of this molecule.
- 12-14** Solve the Hückel equations for a conjugated hydrocarbon chain  $C_8H_{10}$  containing eight carbon atoms. Derive the Hückel eigenvalues and eigenfunctions and the resonance energy of the molecule.



# INDEX

- Ab-initio-ists, 161  
Acetylene molecule, 182  
Ammonia molecule, 181  
Amplitude, 45, 53  
Ångström, 8  
Ångström unit, 8  
Angular frequency, 45  
Angular momentum, 45, 47, 88, 90, 93,  
95, 128, 129, 165  
Anschluss, 20  
Anomalous Zeeman effect, 124, 128,  
157  
Antibonding orbital, 178  
A-posterio-ists, 161  
Aromatic compounds, 183  
Aromatic ring systems, 185, 186  
Atomic number, 142  
Atomic valence orbitals, 176  
Aufbau principle, 142, 143  
  
Balmer, 8  
  
Benzene molecule, 183, 184, 187, 188  
Black-body radiation, 3  
Bohr, 2, 3, 9, 10, 11, 22, 126, 131  
Bohr radius, 100, 102  
Boltzmann, 5, 128  
Born, 2, 3, 11, 12, 13, 16, 20, 21, 22, 56,  
57, 161  
Born-Oppenheimer approximation, 161,  
163, 164  
de Broglie (Louis) 2, 14, 15, 16, 17, 18,  
22, 56, 57, 161  
de Broglie (Maurice) 15, 16  
de Broglie relation, 15, 54, 64  
de Broglie wave, 16, 55, 62, 64  
  
Cadmium atom, 124  
Cartesian coordinates, 46  
Center of gravity, 98, 63  
Chemical bond, 176  
Commutation relations, 90  
Commutator, 89

- Commuting operators, 89  
 Complete sets, 111  
 Compton, 15  
 Compton effect, 15  
 Confluent hypergeometric function, 25  
 Conjugated hydrocarbon chain, 185, 186  
 Copenhagen, 10  
 Correspondence principle, 10  
 Coster, 126  
 Coulomb integral, 137, 148  
 Coulson, 161  
 Courant, 2, 3, 11, 23  
  
 Davisson, 16  
 Debije, 2, 3, 7, 19, 24  
 Determinant, 31, 32, 33, 34  
 Diatomic molecule, 162  
 Differential equation, 24  
 Dirac, 2, 21, 22, 129, 160  
 Directed valence orbital, 178  
 Dispersion relation, 59  
 Dissociation energy, 172, 175, 176  
 Doublet structure, 128, 157  
 Dulong, 7  
  
 Ehrenfest, 2, 3, 10, 125, 128  
 Eigenfunction, 65  
 Eigenvalue, 19, 36, 65, 113  
 Eigenvector, 36, 113  
 Eigenwert, 19  
 Einstein, 2, 6, 7, 16, 125, 131  
 Electron, 9  
 Electron spin, 21, 122, 129, 130, 131  
 Electron spin resonance, 132  
 Elliptical coordinates, 171  
 Elsasser, 16  
 Emission spectrum, 8  
 Ethylene molecule, 181  
 Euler polynomial, 91, 92, 95, 103  
 Excentricity, 48  
 Exchange integral, 137, 148  
 Exclusion principle, 22, 122, 126, 132, 133, 135, 142  
 Expectation value, 57  
  
 Franck, 16  
 Frequency, 4, 11, 53  
 Fundamental law, 39  
 Fourier analysis, 55  
 Fourier integral theorem, 55  
  
 Gallilei, 39  
 Gaussian function, 60  
 Gaussian Program Package, 120, 154  
 Gerlach, 123, 125  
 Goudsmit, 2, 3, 21, 122, 127, 128, 129, 131  
 Gradient, 42  
 Group velocity, 59  
 Gyromagnetic ratio, 129, 131  
 Gyroscope, 24  
  
 Hamilton, 43  
 Hamilton equations of motion, 44  
 Hamiltonian, 12, 43  
 Hamiltonian mechanics, 43  
 Hamiltonian operator, 67, 68  
 Harmonic oscillator, 44, 81, 168  
 Hartree-Fock equation, 149, 150  
 Hartree-Fock method, 120, 135, 146, 175  
 Hartree-Fock operator, 150, 151  
 Heisenberg, 2, 3, 11, 12, 13, 14, 22, 24, 52, 56, 60, 86  
 Hermitian matrix, 29  
 Helium atom, 135  
 Helium atom orbitals, 138  
 Hermitian operator, 66  
 Hexatriene molecule, 188  
 Hilbert, 2, 3, 11, 13, 23  
 Homogeneous polynomial, 91  
 Hückel, 161, 183  
 Hückel method, 183  
 Hund, 143  
 Hund's rule, 143, 155  
 Huygens, 15  
 Hybridized orbital, 178  
 Hydrogen atom, 8, 10, 19, 47, 88, 98, 117, 142

- Hydrogen atom eigenvalues, 103, 104, 105  
Hydrogen molecular ion, 170, 172  
Hydrogen molecule, 170, 173, 176
- Identity matrix, 12  
Indeterminacy principle, 13, 52  
Inert mass, 40  
Integral transform, 55  
Internuclear distance, 172, 175, 176
- Jordan, 11, 12
- K lines, 123, 126  
K shell, 126, 127  
Kepler, 46  
Kepler problem, 46, 88  
Kepler's second law, 48  
Klein, 3, 24  
Kohn, 161  
Koopmans, 151  
Kramers, 2, 3, 10, 11, 152  
Kronecker symbol, 29, 112  
Kummer, 25  
Kummer's function, 25, 26, 27, 83, 100  
Kummer's relation, 27  
Kunsmann, 16
- L lines, 123, 126  
L shell, 123, 126  
Landé, 128  
Laplace operator, 17, 49, 50, 90, 162  
Leibnitz, 32  
Lenard, 6  
Lewis, 6  
Light wave, 4  
Linear algebra, 11, 24  
Linear equations, 35, 36, 112  
Lone pair orbital, 178  
Lorentz, 124, 128
- M lines, 123, 126  
M shell, 123, 126  
Mach, 127  
Matrix, 27  
Matrix multiplication, 28, 29  
Matrix mechanics, 11, 12, 86  
Methane molecule, 179  
Minor, 33  
Molecular-orbital model, 173, 175  
Momentum, 12, 43  
Monochromatic light, 4  
Moseley, 123
- Newton, 1, 6, 23, 39  
Newtonian mechanics, 1  
Nitrogen molecule, 177, 183
- Operator, 66  
Oppenheimer, 161  
Orbital, 135  
Orbital energy, 143  
Oseen, 24  
Overlap integral, 177  
Oxygen molecule, 144
- Particle in a box, 68  
Particle in a finite box, 74  
Pauli, 2, 3, 21, 24, 122, 126, 127, 129, 131, 132, 133  
Period, 4, 48, 53  
Permutation, 30, 133, 136, 147  
Perturbation theory, 108, 113, 114, 115, 119  
Petit, 7  
Phase factor, 53  
Phase velocity, 59  
Photoelectric effect, 6, 7  
Planck, 2, 22  
Planck's constant, 5  
Plane wave, 53  
Polar coordinates, 46, 49, 94, 99, 165  
Polyatomic molecule, 179  
Pople, 161  
Potential curve, 166, 169

- Probability density, 21, 56
- Quantization, 3, 10, 128
- Quantum chemistry, 160
- Quantum number, 101, 102, 126, 127, 132, 135, 142
- Rayleigh, 5
- Rayleigh-Schrödinger perturbation theory, 113
- Reduced mass, 98, 102, 163
- Resonance energy, 188
- Resonance principle, 170
- Resonance structures, 174, 175
- Rigid rotor, 91
- Ritz, 8
- Röntgen, 123
- Rotational constant, 168, 169
- Russell-Saunders coupling, 155, 156, 157
- Rutherford, 9, 123
- Rutherford model, 9
- Rydberg, 8
- Rydberg constant, 8
- Scalar product, 41
- Schrödinger, 2, 14, 17, 18, 19, 20, 22, 98, 114
- Schrödinger equation, 17, 55, 65, 99
- Self Consistent Field method, 120, 146, 150, 151
- Shielding, 152, 153
- Singlet state, 134, 145
- Slater, 145
- Slater determinant, 145
- Slater's rules, 153, 154
- Slater type orbitals, 153
- Sommerfeld, 2, 3, 10, 11, 17, 23, 125, 127, 128
- sp hybridization, 182
- sp<sup>2</sup> hybridization, 182, 184
- sp<sup>3</sup> hybridization, 181
- Special functions, 25
- Specific heat, 7
- Spin functions, 130, 134
- Spin-orbit coupling, 155, 158
- Stark, 117
- Stark effect, 117, 118
- Stationary state, 9
- Stern, 123, 125, 127
- Stoner, 126, 127
- Sturm-Liouville problem, 19
- Tetrahedral geometry, 180
- Thomas, 21, 131
- Thomson, 6, 9
- Transition moment, 11
- Triplet state, 134, 145
- Tunneling, 78
- Uhlenbeck, 2, 3, 21, 122, 127, 128, 129, 131
- Uncertainty principle, 13, 22
- Unit matrix, 29
- Unitary matrix, 30
- Valence-bond model, 173, 175
- Variational principle, 108, 109, 110, 169, 184
- Vector, 40, 41, 42
- Vector model, 155
- Vector product, 41, 46
- Vibrational frequency, 168, 169
- Water molecule, 181
- Wave, 4, 53
- Wavelength, 4, 53
- Wave number, 53
- Wave packet, 58
- Wien, 5
- X rays, 15, 123
- Zeeman, 124, 126
- Zeeman effect, 21, 123, 124, 128
- Zero-point energy, 85