

THE GREENE & STODDARD
FIRE AND LIFE INSURANCE

INTERNATIONAL SERIES IN PURE AND APPLIED PHYSICS

LEONARD I. SCHIFF, *Consulting Editor*

ATOMIC PHYSICS

An Atomic Description of Physical Phenomena

INTERNATIONAL SERIES IN PURE AND APPLIED PHYSICS

LEONARD I. SCHIFF, *Consulting Editor*

- Becker* Introduction to Theoretical Mechanics
Clark Applied X-rays
Edwards Analytic and Vector Mechanics
Finkelburg Atomic Physics
Green Nuclear Physics
Gurney Introduction to Statistical Mechanics
Hall Introduction to Electron Microscopy
Hardy and Perrin The Principles of Optics
Harnwell Electricity and Electromagnetism
Harnwell and Livingood Experimental Atomic Physics
✓ *Harnwell and Stephens* Atomic Physics
Houston Principles of Mathematical Physics
Houston Principles of Quantum Mechanics
Hughes and DuBridge Photoelectric Phenomena
Hund High-frequency Measurements
Kemble The Fundamental Principles of Quantum Mechanics
Kennard Kinetic Theory of Gases
Marshak Meson Physics
Morse Vibration and Sound
Morse and Feshbach Methods of Theoretical Physics
Muskat Physical Principles of Oil Production
Read Dislocations in Crystals
✓ *Richtmyer and Kennard* Introduction to Modern Physics
Schiff Quantum Mechanics
Seitz The Modern Theory of Solids
Slater Introduction to Chemical Physics
Slater Microwave Transmission
✓ *Slater* Quantum Theory of Matter
Slater and Frank Electromagnetism
Slater and Frank Introduction to Theoretical Physics
Slater and Frank Mechanics
Smythe Static and Dynamic Electricity
Squire Low Temperature Physics
Stratton Electromagnetic Theory
Thorndike Mesons: A Summary of Experimental Facts
Townes and Schawlow Microwave Spectroscopy
White Introduction to Atomic Spectra
-

The late F. K. Richtmyer was Consulting Editor of the series from its inception in 1929 to his death in 1939. Lee A. DuBridge was Consulting Editor from 1939 to 1946, and G. P. Harnwell from 1947 to 1954.

ATOMIC PHYSICS

An Atomic Description of Physical Phenomena

GAYLORD P. HARNWELL

Professor of Physics, University of Pennsylvania

WILLIAM E. STEPHENS

Professor of Physics, University of Pennsylvania

McGRAW-HILL BOOK COMPANY, INC.

New York Toronto London

1955

ATOMIC PHYSICS

Copyright © 1955 by the McGraw-Hill Book Company, Inc. Printed in the United States of America. All rights reserved. This book, or parts thereof, may not be reproduced in any form without permission of the publishers.

Library of Congress Catalog Card Number 54-10635

THE MAPLE PRESS COMPANY, YORK, PA.

To Mollie and Helen

PREFACE

The text that follows represents an effort to present to students in their first year of graduate study the essential physical ideas upon which the current atomic theory of matter rests. It has evolved during the period from 1947 to 1954 in connection with the course given by the authors at the University of Pennsylvania. The students entering their first year of graduate work, during which this course is commonly taken, have all nominally had undergraduate courses in general and atomic physics, and also in mechanics and electricity and magnetism. However, their backgrounds and the quality of their preparation are generally very diverse, and, as indicated by the material in the first chapter, several weeks are given to a general summary of those aspects of classical physics which are essential to the concepts upon which atomic theory is based and to those fields of classical experimentation from which the techniques of atomic investigations have evolved. As the students are most commonly familiar with classical physics in terms of the absolute practical, or mks, system of units, this system of units has been used in the course and adopted in the text, though somewhat of a break with precedent in the field of atomic physics is thereby involved.

Atomic physics has acquired such an extensive history and so vast a literature during the first half of the twentieth century that it is no longer feasible to adopt a strictly historical type of approach to a one-year course, much less to make any pretense at completeness of coverage. The early forms of the quantum theory and the early work on gas discharges, X rays, etc., receive but cursory mention except as background against which the atomic concepts took form. Also the more peripheral areas which are covered by special courses such as quantum mechanics, complex spectra, polyatomic molecules, crystallography, nuclear structure, etc., are touched on very lightly if at all because of both lack of space and lack of any special competence on the part of the authors. The emphasis of the book is upon the extension of the basic classical concepts of physics into the realm of atomic phenomena and on the evolution of those particularly central and elemental quantum concepts, having no classical counterpart, which uniquely characterize the physics of elementary particles.

The last two chapters are then concerned with the statistical description of these particles and their aggregates; in particular those properties of gases, liquids, and solids with which the current theory is adequate to deal on an elementary mathematical level. Throughout the book an attempt is made to focus on the physics involved rather than upon the mathematical apparatus and techniques. The authors are well aware that they have fallen somewhat short of their objectives in this regard and have probably erred in some places by being too detailed and explicit and in others by being too brief or scanty. The general tenor of the presentation will probably be found more congenial to those experimentally minded physicists who, like the authors, admire mathematical elegance but are unable to practice it themselves.

The authors are greatly indebted to their colleagues at the University of Pennsylvania for their continued interest, their helpful suggestions and advice, and their constructive criticism of many sections. They are particularly indebted to Professors Herbert Callen, Sherman Frankel, Vernon Hughes, Peter Landsberg, Enos Witmer, and Paul Yergin for their assistance in the final stages of preparing this manuscript for the printers.

GAYLORD P. HARNWELL
WILLIAM E. STEPHENS

CONTENTS

PREFACE	vii
1. CLASSICAL FOUNDATIONS OF ATOMIC THEORY	1
1.1. Scope and Methods of Physical Science	1
1.2. Evolution of Atomic Concepts	3
1.3. Newtonian Mechanics	6
1.4. Wave Motion	12
1.5. Thermodynamics	15
1.6. Relativistic Mechanics	18
1.7. Electricity and Electromagnetism	23
1.8. Electromagnetic Waves and Oscillators	28
1.9. Relativistic Formulation of Electrodynamics	34
Problems	36
References	39
2. ATOMIC NATURE OF MATTER AND RADIATION	40
2.1. Early Evidence for Atomicity	40
2.2. Avogadro's Number	42
2.3. Charges and Masses of Atomic Components	50
2.4. Corpuscular Nature of Radiation	60
2.5. Wave Nature of Matter	65
2.6. Angular Momenta and Magnetic Moments of Particles	70
2.7. Higher-order Moments	85
Problems	90
References	93
3. ATOMIC STRUCTURE	95
3.1. Early Spectroscopy	95
3.2. Evidence for the Nuclear Atom	99
3.3. Experimental Determination of Energy Levels	108
3.4. Schrödinger Theory	121
3.5. Quantum Theory of Radiation	135
3.6. Quantum Radiation Processes	143
Problems	154
References	159

4. ELECTRON SPIN AND POLYELECTRONIC ATOMS	160
4.1. Concept of Spin and Simple Magnetic Effects	160
4.2. Consequences of a Relativistic Formulation	166
4.3. One-electron Zeeman Effects	180
4.4. Hyperfine Structure	188
4.5. Identical Particles and the Pauli Principle	197
4.6. Quantitative Methods; Helium	210
Problems	218
References	222
5. MOLECULAR STRUCTURE	224
5.1. Electron and Nuclear Locations	224
5.2. Nature of Molecular Energy Levels	233
5.3. Molecular Electronic States	236
5.4. Molecular Nuclear Motion	249
5.5. Molecular Spectroscopy	255
Problems	275
References	279
6. STATISTICS OF ATOMIC PROCESSES	280
6.1. Statistics of Observation	280
6.2. Thermal Equilibrium	285
6.3. Applications of Classical Statistics	293
6.4. Quantum Statistics	299
6.5. Applications of Quantum Statistics	304
Problems	317
References	321
7. ELEMENTARY PROPERTIES OF MATTER	322
7.1. Classical Properties of Matter	322
7.2. Perfect Gases	324
7.3. Rate and Transfer Processes	335
7.4. Imperfect Gases	342
7.5. Approach to the Liquid State	349
7.6. Mechanical and Thermal Properties of Crystals	357
7.7. Electron Band Theory of Crystals	364
Problems	382
References	387
RELATIONSHIPS BETWEEN UNITS	389
PRINCIPAL ATOMIC CONSTANTS	390
INDEX	391

ATOMIC PHYSICS

An Atomic Description of Physical Phenomena

CHAPTER 1

CLASSICAL FOUNDATIONS OF ATOMIC THEORY

1.1. Scope and Methods of Physical Science. The study of physical science is a variety of intellectual exploration. Of necessity one progresses from the known to the unknown erecting a framework of theory within which quantitative observations form an ordered pattern. This process is accomplished in terms of mental concepts brought to the task, and these are themselves derived from earlier and simpler sensory experiences. Of course many of the primary facts of observation are of great interest in themselves. But to a still greater degree their interrelationships and their implications contribute to our intellectual grasp of the totality of phenomena that constitutes our environment. The objective of the process in which our mental abilities are employed to digest the vast diversity of our observations is the formulation in words or other symbols of an adequate and ordered description of all physical phenomena in terms simple enough for us to grasp. Such a description is called a scientific theory. The word "understanding" is generally used in connection with observed phenomena in the sense of the achievement of an adequate description or the formulation of a satisfactory theory. Understanding is also used in reference to those irreducible concepts in terms of which a theory is formulated, as indicating a mental comprehension or intuitive grasp of these concepts and their relationships with one another.

A description or theory is considered adequate to the extent that it provides a universal, precise, and simple account of observation. Present theories are all inadequate in one or more of these respects, and they do not provide a simple ordering of all of the essential aspects of observation. The objective of science is to extend our observations and to improve the theories that order them. A theory is conceived and evolves by the interplay of inductive and deductive processes. Initial simple observations and imaginative triumphs of conception suggest hypotheses. Further observation and conceptual syntheses contradict, confirm, modify, or extend these hypotheses. Additional concepts are evolved, and relationships and suggestive analogies are noted which indicate profitable directions for further observation. Simplification and unifica-

tion are striven for, and irrelevancies and errors are sought out and eliminated. The raw data of sensory perception are thus eventually ordered by the resulting description. Our understanding deepens through a simplification of the elementary essential concepts that remain. Uniqueness can hardly be considered as an objective for there are no adequate criteria with which to assess the possibility of constructing alternative explanations or descriptions. But as a theory develops the breadth to account for a wide diversity of observations and as the concepts it employs become reduced to the simplest and most elemental terms, the possibility of formulating completely different alternative descriptions appears to become more and more remote.

A scientific theory resembles a fabric in that its strength and acceptability come from the manifold of separate interwoven threads of connection between observations of different types. The greater the intimacy and extent of relatedness provided by a theory between the phenomena of observation, the more significant and important does the theory become. A description that brings together but a few phenomena in terms of simpler common concepts is a suggestive beginning warranting further exploration, but it is a web of only a few strands upon which little reliance is placed. As the number and variety of observations compatible with the initial small number of hypotheses are extended, the theory takes on generality, it represents a more obvious economy of thought, and it eventually appears actually to provide in some sense an intellectual grasp of a whole aggregate of separate observations and phenomena. The interlacing of the strands of connection between observations is so compact that additional confirming observations are immediately suggested, and the pattern that is finally presented takes on a significance of its own, transcending the sum of the individual observations from which it has been developed. Certain portions of this fabric even tend to take on a rigidity which conditions our acceptance of new phenomena. Thus the basic position of the conservation laws, such as those of mass, energy, and momentum, in dynamical theory imposes certain habits of thought in terms of which the phenomena presented by new elementary particles, such as the varieties of mesons currently being observed, are interpreted. The present difficulties which are being encountered in achieving a consistent theory of these particles may necessitate a restatement of fundamental laws in broader terms than are known at present.

The science of physics relates primarily to the basic phenomena which are associated with the central concepts of mass, charge, motion, and energy. In the previous century physics was considered to represent a more or less unrelated group of fields such as mechanics, properties of matter, heat, sound, light, electricity, and magnetism. The atomic concept of matter and the observed interactions of matter and radiation supplied

a common basis for unification of these separate fields and also led to a great expansion of the phenomena included within the field of physics. The experimental observations and the theory which unites them are in a continuous state of flux and growth. The present concepts do not appear to be altogether adequate for the construction of a completely satisfactory physical theory, and in spite of the great progress that has been made our understanding of elementary particles is but fragmentary and imperfect. Much of a very fundamental nature remains to be done, and, in consequence, a grasp of the general methodology and of the criteria for the attainment of the objectives of science is particularly important.

1.2. Evolution of Atomic Concepts. The growth of physical theory is a cumulative process which takes place by painstaking accretion with occasional brilliant flashes of insight, proceeding essentially from the known to the unknown. The concepts and methods of thought that have gone into the construction of atomic theory are those which were current in classical physics together with those drawn from fields of mathematics and certain other quite novel concepts that have been slowly and painfully evolved in the course of recent physical research. Classical physics dealt almost exclusively with objects of simple sensory perception. It was concerned with "man-sized" objects and "man-sized" forces. Color, sound, taste, smell, and surface texture were significant attributes of the objects of observation. Concepts of wave motion came from observing taut strings and the transient conformation of a liquid surface. The apparent permanence of most material objects suggested the reasonableness of conservation laws. The first notions of force and weight and the properties of fluids came from muscular and tactile sensations. These types of qualitative observation were very important in the development of the background of mental images in terms of which descriptions could be formulated and diverse observations related to one another.

Physics, however, is characterized by its emphasis on quantitative relationships. Careful observations led to a knowledge of the existence of such relationships, and mathematics provided a simple and unambiguous method of stating them. The type of sensory data which could yield precise quantitative information assumed the major role in contributing to the development of the science. The comparison of lengths, the recording of pointer readings of clocks and balances, and the use of spring-type standards of forces and torques which produced other pointer readings led to the primary reliance on this simple use of the sense of sight as the chief source of quantitative information. The emphasis on numerical description and accuracy tended to sharpen the basic concepts and to reduce the importance of those which could not be stated with

precision. The mathematical techniques which were used to express the relationships between concepts were found to be applicable widely throughout the classical subfields of physics and in fact formed the principal unifying link between them. The ideas of coordinate systems and their transformations, vectors, tensors, invariants, and many other concepts from the field of pure mathematics contributed immeasurably to the evolution of simple, precise formalisms in terms of which phenomena could be described and theories expressed.

As the concept of the atomicity of matter evolved from the first simple chemical and physical observations it became clear that the ultimate particles were much smaller than the samples of matter previously dealt with. If these ultimate particles are too small to be seen or felt, localization can only be inferred. Hearing is of no use as a sense for direct perception, and color also ceases to be a significant attribute. Thus it is seen that only certain of the concepts of classical physics, primarily those of a precisely defined and quantitative character, are capable of meaningful extension down to these minute dimensions and hence likely to be of use in the characterization of atoms. The common impressionistic descriptions of everyday objects and occurrences in terms of shape, color, size, texture, and sound find no application in atomic descriptions. Ordinary language largely fails as a method of precise characterization; few common adjectives and but a handful of verbs remain useful. Of course the language of common communication is used to delineate those concepts which are found capable of meaningful extension into the atomic domain, but the precise statements that are made regarding their relationships must be couched in mathematical terms. The field is one of stark utilitarian simplicity; atoms and their constituents have few properties, so there is little variety in what can be said about them. Nevertheless, the statements that can be made are highly precise and carry as their implications all the more complex descriptions of large-scale phenomena.

The attributes with which present theories endow atomic particles may well be only the simplest of those which they will eventually be found to possess. Already, novel concepts such as *parity* and *spin*, which have no direct large-scale classical analogues, have had to be devised. Others may yet be required as the theory evolves. The relationship between direct experiment and ultimate theory in atomic physics is a more difficult and derivative one than in the case of gross or large-scale classical physics. Individual atomic events are not observed directly, though their consequences may be, and hence greater reliance must be placed on logical inference and methodology than in most other branches of science. The relation between an atomic property and an observational test is often a complex one and requires considerable intervening logical

connection. This restricts the use of simple, direct inference, and the results of many lines of inquiry are frequently required to establish conclusively the existence of an atomic characteristic.

The concepts of atomic physics are drawn from classical physics, and the instruments used in atomic research are those which were evolved from classical laboratory techniques. Thus an intimate knowledge of the fields of mechanics and electricity is essential to an understanding of atomic theory. Many of the experimental verifications of atomic properties depend on observations of the properties of matter on a large scale. Thus the study of thermodynamics and the statistical behavior of large numbers of atoms, which provides a connection between atomic attributes and the gross properties of matter, is very important. A knowledge of the basic principles in these classical fields and a familiarity with mathematical techniques concerned with them are necessary before undertaking the study of atoms, their elementary constituents, and their combinations in molecules and crystals. The extensive success that has been achieved in accounting for the properties of atoms and atomic systems constitutes strong evidence for the general adequacy of present concepts and theories. The concepts and the theory are basically simple, though somewhat unfamiliar, and with a remarkable elegance and economy of hypotheses they account for a vast diversity of phenomena throughout all of physical observation.

The theory, however, is by no means perfect, and its further development constitutes one of the most absorbingly interesting fields of scientific activity today. One class of current problems centers on the technical difficulties surrounding the mathematical description of complex atoms and molecules involving many electrons and nuclei. Here the laws and concepts appear to be sufficiently well understood to assure a detailed and satisfactory account if the complicated mathematical expressions could be handled. The second and more intriguing class of problems is that in which ideological difficulties are clearly fundamental. These involve the nature of elementary particles which are found to be either permanently or transiently associated with nuclear and high-energy processes. Here present concepts may well be inadequate for construction of a satisfactory theory; certainly more extensive experimental and intellectual investigation will be necessary before the interaction between elemental entities can be understood. These are the frontiers of atomic physics which are now the focus of research effort.

The sections which follow in this chapter summarize those classical concepts and procedures which are basic to the formulation of the atomic theory developed in subsequent chapters. The purpose is merely to organize and recall those fields of mechanics, thermodynamics, and electricity which are assumed to be familiar from courses in these subjects.

The discussion of electromagnetic dipole radiation is somewhat more extended, as this matter is frequently not treated in elementary courses in electricity. More complete accounts of these subjects are to be found in the general references listed at the end of this chapter.

1.3. Newtonian Mechanics. Newtonian mechanics deals with the commonly encountered motion of material objects. A description of this motion involves a statement of the positions of the objects as functions of time, and for this purpose coordinate systems and standards of *length* and *time* are the first requirements. The position of a point relative to an origin may be given in terms of a cartesian coordinate system and a unit of length. The variation in position of a point or a series of points forming an extended body requires for its quantitative description the introduction of a unit of time. The definition of units is always quite arbitrary, and here the so-called *mks* system will be adopted. In this system the unit of length is the meter, which is the interval between two marks on an arbitrarily chosen metal bar, and the unit of time is the second, which is an arbitrarily chosen fraction of the mean solar day. Though the choice of units is inherently arbitrary and quite without effect on the content of physical laws, the constancy of the velocity of light and the characteristic frequencies of the electromagnetic radiation emitted by atoms suggest the choice of certain other standards of length and time as having a more natural relationship to invariant quantities of observation. Thus the wavelength of a line in the red portion of the cadmium spectrum or a spectrum line observed to be emitted from a mercury glow discharge when the mercury consists of a single isotope is more readily reproducible than the arbitrary meter and can be defined and compared with a precision equal to or even greater than that with which meter sticks can be compared. Similarly, the reciprocal of natural atomic or molecular frequencies can be used as standards of time in "atomic clocks" such as those of the U.S. National Bureau of Standards or the British National Physical Laboratory. Such standards appear to be as reliable, reproducible, and convenient as the one derived from astronomical observation, and very high accuracies of comparison can be achieved with ease.

The initial and final data in a mechanical problem, which is the prototype of any set of quantitative observations, are the positions of objects at the initial and final times. To describe the phenomena that are observed, additional concepts are convenient. Two of these are the derived kinematic concepts of *velocity* and *acceleration*, which are the rate of change of position and rate of change of velocity, respectively. The laws of classical, or Newtonian, mechanics relate the motion of bodies or certain readily identifiable quantities of matter to what are called *forces*. The concept of force is one of which we have an immediate,

though of course only qualitative, kinesthetic appreciation. The phrase "quantity of matter" is made precise by introducing the concept of *mass*. The mass of a body is a quantity that can be measured, most commonly by the use of a gravity balance. The unit of mass in the system here adopted is the kilogram, which is the mass of an arbitrarily chosen block of platinum approximately equivalent in mass to a cube of water a tenth of a meter on a side. It is permanent and reproducible and can be compared with other masses to about the same accuracy as for the other two basic units. Here again an atomic standard may eventually offer certain advantages, but presently available techniques for comparing atomic masses with gross masses do not as yet offer acceptable precision.

The laws of mechanics can be stated most generally and conveniently in terms of the concept of *momentum*, which is defined as the product of mass and velocity. Mass is a scalar quantity, generally designated by the letter *m*, and velocity (like position relative to an origin) is a vector quantity having both magnitude and direction, generally designated by the letter *v*. Thus the momentum *p* is defined as $\mathbf{p} = m\mathbf{v}$. The fundamental equation of mechanics which relates mass, motion, and force is the following:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} \quad (1.1)$$

The quantity on the right is the rate at which momentum changes with time, and this equation defines the mechanical concept of force. It also defines the unit of force in the system of units used for measuring momentum and time. In the mks system the unit of force which causes a change of momentum at the rate of one kilogram meter per second per second is called the newton. Equation (1.1) is a vector relation illustrated in Fig. 1.1. There are three analogous equations between the vector components of the forces and rates at which the vector components of the momentum change. This definition of force is observed to be in general qualitative concordance with our kinesthetic understanding of the term.

Through this equation, physics deals with forces of all types which are thought of in various connections as having different natures or origins. Classical mechanics is concerned to a great extent with contact forces such as those involved when one pushes against an object or when two pieces of matter collide. In the description of such events surfaces and areas of contact are involved, but such concepts are found to be of little use in increasing our understanding of forces on an atomic scale. Mechanics is not concerned with the nature of the forces involved since

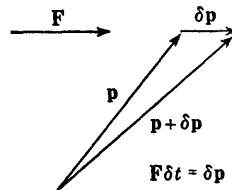


FIG. 1.1. Fundamental law of mechanics.

these do not appear in the final answer to a problem. The initial data are the positions of certain objects at certain times, and the predictions of mechanics relate to their positions at later times. Forces are eliminated from the final result by the intervening calculations. Physics as a whole, however, is concerned with the nature of the forces that are observed to bring about changes in momenta, and to the extent that the concept of force is utilized the objective is to account for the presence of forces in terms of simpler and more fundamental relationships between interacting particles.

Historically, the first reduction of a force into more elemental terms was the formulation of the universal law of gravitation by Newton. This law states that a property of a mass is that it exerts on all other masses an attractive force which is proportional to the masses involved and inversely proportional to the square of their separations. This may be expressed quantitatively in terms of kilograms and meters by

$$\mathbf{F}_{m_1} = G \frac{m_1 m_2}{r^2} \mathbf{r}_{12} \quad (1.2)$$

where \mathbf{F}_{m_1} is the force in newtons on the mass m_1 exerted by the mass m_2 at a distance r , \mathbf{r}_{12} is a unit vector from m_1 toward m_2 , and G is the gravitational constant determined experimentally to be $(6.670 \pm 0.005) \times 10^{-11}$ newton-m² kg⁻². The equivalence between inertial mass, which is the m of the equation $\mathbf{p} = m\mathbf{v}$, and gravitational mass, which is the m of Eq. (1.2), is shown by the concordance of Eqs. (1.1) and (1.2) with experimental observation. A second elemental source of force, which is particularly important in atomic physics, is that which is identified with electric charges. Mechanics per se, however, is not concerned with these matters but rather with the motion of particles when subjected to forces of whatever type.

Equation (1.1) is incomplete as a statement of the principles of mechanics since forces do not appear in the final answer to a mechanical problem. An observation, also due to Newton, relating the forces between two interacting bodies must also be included. This can be stated as follows: The force exerted by one body on another is equal in magnitude and opposite in direction to that exerted by the second body on the first. Alternatively, this may be stated as the *conservation of momentum* by saying that the total momentum of a closed system of particles subject only to forces between them and not to forces of external origin is constant. The equivalence of these statements can be seen by the use of Eq. (1.1). If, within a large group of interacting particles, \mathbf{F}_{ij} is the force exerted by particle i on particle j at the time t and \mathbf{p}_j is the momentum of the j th particle, Eq. (1.1) for all the particles may be added together to yield

$$\sum_i \sum_j \mathbf{F}_{ij} = \frac{d}{dt} \sum_j \mathbf{p}_j \quad (i \neq j)$$

If $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$, which states the equality of action and reaction, the sum on the left cancels in pairs and the rate of change of $\sum_j \mathbf{p}_j$ is zero, or $\sum_j \mathbf{p}_j$ is a constant. The law of conservation of momentum is a most fundamental one and is applicable in this form to all simple mechanical phenomena. In the case of electromagnetic phenomena all the forces acting between the elementary constituent particles are not of such a type that their directions coincide with the lines of centers and the mechanical momentum is not conserved. However, following the mechanical tradition, it has been found possible to formulate the theory of electromagnetic phenomena in such a way that an analogous generalization emerges, and what may be considered as the sum of the mechanical and electromagnetic momenta of a complete system of interacting particles is an invariant quantity.

Considering only the equal-and-opposite line-of-center type of force, a quantity known as the *angular momentum*, defined as the vector product of the vector representing distance from some arbitrary point of origin and the vector momentum, displays a conservation property. The total angular momentum \mathbf{L} of a large group of particles is given by the vector sum of the angular momenta of the individual particles,

$$\mathbf{L} = \sum_i \mathbf{l}_i = \sum_i m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i$$

Taking the derivative of \mathbf{L} with respect to the time,

$$\frac{d\mathbf{L}}{dt} = \sum_i m_i \dot{\mathbf{r}}_i \times \dot{\mathbf{r}}_i + \sum_i m_i \mathbf{r}_i \times \ddot{\mathbf{r}}_i$$

The vector product of $\dot{\mathbf{r}}_i$ with itself vanishes, and by Eq. (1.1) $m_i \dot{\mathbf{r}}_i$ can be written as $\mathbf{F}_i + \sum_j \mathbf{F}_{ji}$, where \mathbf{F}_i is any external force acting on particle i and \mathbf{F}_{ji} is the force exerted on particle i by particle j . Thus

$$\frac{d\mathbf{L}}{dt} = \sum_i \mathbf{r}_i \times \mathbf{F}_i + \sum_i \sum_j \mathbf{r}_i \times \mathbf{F}_{ji} \quad (i \neq j)$$

Since $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$ the double sum can be written in pairs as

$$\frac{1}{2} \sum_j \sum_i (\mathbf{r}_j - \mathbf{r}_i) \times \mathbf{F}_{ji}$$

If $(\mathbf{r}_j - \mathbf{r}_i)$, which is the vector from i to j (see Fig. 1.2), is collinear with \mathbf{F}_{ij} , this sum vanishes and the internal forces make no change in total angular momentum of the system of particles. The sum $\sum_i \mathbf{r}_i \times \mathbf{F}_i$ is defined as the total torque \mathbf{T} , and the equation of rotational motion is

$$\mathbf{T} = \frac{d\mathbf{L}}{dt} \quad (1.3)$$

If there are no external forces, $\mathbf{F}_i = 0$; hence $\mathbf{T} = 0$, and \mathbf{L} does not change with time. Thus the angular momentum is said to be conserved.

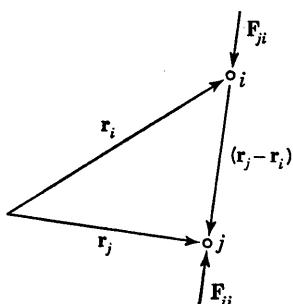


FIG. 1.2. Forces collinear with vector separations.

The principle of conservation of angular momentum appears to be as fundamental and capable of as deep extension into the realm of atomic physics as the conservation of linear momentum.

The concept of *energy*, which evolved from that of work, is also of primary importance in atomic theory. Muscular sensations and the relation of human effort to the accomplishment of alterations in the positions of bodies of matter led to the concept of work, and in general conformity with these ideas work is quantitatively defined for physical

purposes as the integral of the scalar product of the force and the displacement which describes the resulting motion. The energy concept has been greatly extended from the form in which it initially appeared in the subject of mechanics, and it is very largely through this extension of the concept of energy that the prevailing unification has been brought about in the subfields of classical physics. The adoption of the principle of the conservation of energy, which survived the extensions and inclusions unifying the classical fields of physics, has had a profound effect upon the concepts of atomic physics and the theory that interrelates these concepts. To kinetic energy and the various forms of potential energy of gravitation, hydrodynamics, and elasticity was added the concept of heat as a form of energy about 100 years ago. The concept of mass as a form of energy introduced at the beginning of this century was a great advance which followed closely on the application of the energy concept to electric, magnetic, and electromagnetic phenomena. More recently, still further extensions are involved in the so-called *exchange energy* of quantum mechanics and the nuclear binding energy.

A mechanical example provides the simplest illustration of the energy concept. If the vector force \mathbf{F} of Eq. (1.1) is multiplied scalarly by the infinitesimal vector $d\mathbf{r}$ and the product integrated from some arbitrary

point 1 to point 2 in space, the result for a constant mass may be written

$$\begin{aligned}\int_1^2 \mathbf{F} \cdot d\mathbf{r} &= \int_1^2 m\ddot{\mathbf{r}} \cdot d\mathbf{r} = \int_1^2 m\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}} dt \\ &= \int_1^2 \frac{d}{dt} \left(\frac{1}{2} m \dot{\mathbf{r}}^2 \right) dt = \frac{1}{2} m v_2^2 - \frac{1}{2} m v_1^2\end{aligned}$$

where \mathbf{v} is written for $\dot{\mathbf{r}}$. The $d\mathbf{p}/dt$ of Eq. (1.1), the negative of which is sometimes called the inertial reaction, has thus associated with it an inertial energy $\frac{1}{2}mv^2$ known as kinetic energy. Assume that the external force \mathbf{F} which is applied to the body is of such a type that its components are the negative spatial partial derivatives of a scalar function V . Then $\mathbf{F} = -\nabla V$, where $\nabla = i \partial/\partial x + j \partial/\partial y + k \partial/\partial z$, and

$$\int_1^2 \mathbf{F} \cdot d\mathbf{r} = - \int_1^2 (\nabla V) \cdot d\mathbf{r} = - \int_1^2 \frac{\partial V}{\partial r} dr = -(V_2 - V_1)$$

Equating these two forms of $\int \mathbf{F} \cdot d\mathbf{r}$,

$$V_2 + \frac{1}{2}mv_2^2 = V_1 + \frac{1}{2}mv_1^2 \quad (1.4)$$

or the sum of the potential energy and the kinetic energy is a constant.

Equation (1.4), which is a statement of the conservation of energy in the simplest form in which only kinetic and potential energies appear, is frequently the most convenient starting point for the solution of mechanical problems. When there are forces of a frictional or viscous type, which are not derivable from a potential-energy function, mechanical energy is transformed into heat. In accordance with the second law of thermodynamics, equilibrium is characterized by the reduction of mechanical energy to a minimum consistent with the external conditions. The useful consequences of this statement form the subject of thermodynamics. Also in the case of electromagnetic forces there is not, in general, an energy function from which the forces may be derived by simply taking its negative gradient. However, if the energy concept is extended to include electromagnetic fields and potentials, a generalized energy function may be constructed and used for the solution of problems.

The most useful way of formulating the laws of mechanics for application in the field of atomic physics is in terms of a total energy function which displays a property of constancy similar to Eq. (1.4). In the case of a considerable number N of interacting mechanical particles for which there is a simple potential-energy function dependent only on their spatial coordinates, Eq. (1.4) may be simply generalized to

$$\sum_{i=1}^N [\frac{1}{2}m_i v_i^2 + V(x_i, y_i, z_i)] = \text{constant}$$

If this is written in terms of the momenta rather than the velocities, it is an example of the so-called *Hamiltonian energy function* H :

$$H = \sum_{i=1}^N \left(\frac{p_i^2}{2m_i} + V \right) = \text{constant} \quad (1.5)$$

Writing q_i as one of the $3N$ cartesian coordinates of the N particles, the equations of motion are seen from Eq. (1.1) to be in this case

$$\dot{p}_i = -\frac{\partial V}{\partial q_i}$$

and the velocities are $\dot{q}_i = \frac{\partial}{\partial p_i} \left(\sum_i \frac{p_i^2}{2m_i} \right)$, or alternatively,

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i}$$

It is shown in treatments of mechanics that the function H and the first-order partial differential equations above derived from it have the same meaning and are of the same form even if cartesian coordinates are not used. Consider that a new set of $3N$ coordinates is introduced by some point transformation such as $x_i = x_i(q_1, q_2, \dots)$ and similarly for the y 's and z 's. Consider also that the momenta conjugate¹ to the new coordinates are defined as the partial derivatives with respect to the q 's of the so-called *Lagrangian function* L , which is the difference between the kinetic and potential energies expressed in terms of the coordinates and their time rates of change. Then $p_i = \partial L(q_i, \dot{q}_i)/\partial \dot{q}_i$, and the Hamiltonian is the function $H = \sum_i \dot{q}_i p_i - L$ expressed in terms of the q 's and the p 's.

The equations of motion may be derived from the Hamiltonian by the same partial differential equations given above. The p 's are then not necessarily linear momenta but, for example, are angular momenta if the q 's are angles. It may even be shown that it is possible to extend the range of application of this method to include certain cases of velocity-dependent forces such as those which arise in electromagnetic problems. Here the momentum conjugate to a given coordinate is not simply the mechanical momentum but includes as well what may be interpreted as electromagnetic momentum.

1.4. Wave Motion. The concept of oscillatory behavior, or wave motion, is a basic one which finds wide application in atomic physics. The fundamental ideas in this field derive largely from experience with

¹ Conjugate in the canonical sense.

vibrating strings or bars or the motion of disturbances along wires or on the surfaces of liquids. The intuitive appreciation of these phenomena must, however, be stripped of all unessential particulars before the ideas are suitable of extension to atomic processes. In contradistinction to the familiar instances such as waves on a pond, where the undulations are a transitory feature of a familiar object, the wave aspect of elementary particles is possibly their most fundamental characteristic, and it is not possible to reduce this concept further by forming any adequate mental picture of what substance, if any, is undergoing pulsation. A wave concept of matter is said to be employed and familiar ideas of frequency and wavelength are used in discussing it simply because of the similarity

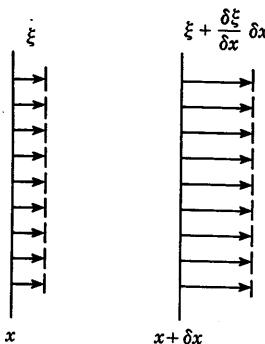


FIG. 1.3. Laminar displacement as a function of the coordinate.

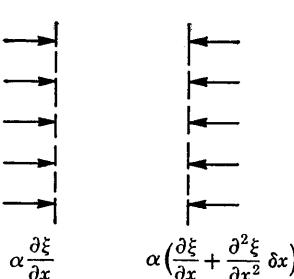


FIG. 1.4. Forces on lamina.

in form between the differential equations and their solutions that pertain to large-scale wave motion and those that provide an adequate description of atomic phenomena.

As a representative instance of wave motion consider a small region of some elastic medium in which displacements from the normal positions of the particles constituting the medium occur in one direction only, say that of the coordinate x . Let ξ represent the displacement of the particles, and let ξ be a continuous function of x . A compression or extension of the material occupying the lamina δx in the neighborhood of x occurs if ξ is not constant (see Fig. 1.3). If the material obeys Hooke's law, there is a force per unit area proportional to the elongation per unit length, i.e., the force per unit area acting across the faces of the lamina is $\alpha \partial \xi / \partial x$, where α is Hooke's constant. The net force per unit area across the two bounding planes of a lamina δx in thickness is thus $\alpha (\partial^2 \xi / \partial x^2) \delta x$ (see Fig. 1.4). By Eq. (1.1) this must equal the inertial reaction of the material in δx , which, if the density is ρ , is given by $\rho \delta x \partial^2 \xi / \partial t^2$. Thus $\alpha \partial^2 \xi / \partial x^2 = \rho \partial^2 \xi / \partial t^2$, or

$$\frac{\partial^2 \xi}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 \xi}{\partial t^2} \quad \text{where } v = \left(\frac{\alpha}{\rho}\right)^{\frac{1}{2}}$$

This type of differential equation characterizes what is called wave motion because, as can be seen, any function of the form $\xi(x \pm vt)$ satisfies it. If the displacement ξ has some particular value at x_1 and t_1 , it will have this same value at a neighboring point x_2 at a subsequent time t_2 if

$$x_1 \pm vt_1 = x_2 \pm vt_2 \quad \text{or} \quad \frac{x_2 - x_1}{t_2 - t_1} = \pm v$$

Hence v represents the velocity of propagation of the displacement in the sense that v , equaling the ratio of the spatial to the temporal interval, would represent the motion of a disturbance from point 1 to point 2.

$$k = \frac{\omega}{v} = \frac{2\pi}{\lambda}$$

The wave equation can be generalized to three dimensions by an obvious extension of the above argument. If

$$\nabla^2 \xi = \frac{\partial^2 \xi}{\partial x^2} + \frac{\partial^2 \xi}{\partial y^2} + \frac{\partial^2 \xi}{\partial z^2}$$

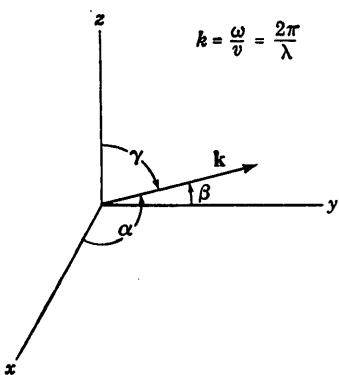


FIG. 1.5. Propagation vector \mathbf{k} in the direction of the wave normal.

in the direction of the vector \mathbf{k} (equal in magnitude to ω/v) is (Fig. 1.5)

$$\xi = C e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$$

as can be seen by substitution. By the use of Fourier's theorem an arbitrary disturbance can be built up by combining a series of such waves of different angular frequencies. The total disturbance at $r = 0$ would then be

$$f(t) = \sum_{s=-\infty}^{s=\infty} C_s e^{is\omega t}$$

where

$$C_s = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} f(t) e^{-is\omega t} dt$$

The function $f(t)$ is periodic, reproducing itself in each interval of length $2\pi/\omega$. If it is desired to expand a nonperiodic function, it may be shown, subject to certain limitations, that

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\omega) e^{i\omega t} d\omega$$

$$g(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

The functions $f(t)$ and $g(\omega)$ are known as a pair of Fourier transforms.

1.5. Thermodynamics. Thermodynamics is the extension of mechanics to include *heat* as a form of energy and *temperature* as a parameter for specifying the state of a system. This extension represents a very convenient formulation of the results of experimental observation and is widely used throughout the classical fields of physics and chemistry. This formulation is not particularly appropriate to atomic physics, where the introduction of statistical methods and concepts leads to a description of the distribution of energy among the elementary particles that constitute the system. Nevertheless, the concepts of thermodynamics retain their usefulness as parameters in such a statistical formulation, and the nomenclature of thermodynamics continues to be a most useful one for the description of large-scale phenomena.

The first law of thermodynamics is simply the statement in differential terms of the conservation of energy with the inclusion of heat as a form of energy. This is in accordance with the general experience of thermal processes in which a body becomes hotter as work is expended upon it in a forging or rolling operation. Quantitatively it is not heat that is measured but a parameter that is called temperature. Like force, it is a quantity of which we have a qualitative sensory appreciation. The concept of temperature evolved from the gas laws, and a gas thermometer provides one of the most precise methods of measuring it. The second law of thermodynamics states that, when two bodies initially at different temperatures are placed in thermal contact with one another, heat flows or energy is transferred from the hotter to the colder body (entropy increases).

The conservation of energy is expressed most usefully in a differential form. Let a certain quantity of heat δQ be generated, say by the combustion of fuel within a container, and at the same time, possibly by the expansion of a piston against a force, let a certain amount of work δW be done. The symbol δU represents a possible change in the internal energy of the system, which is presumably something that could be expressed in detail in terms of coordinates if it were possible to describe the position and the motion of each of the particles making up the sys-

tem. The conservation of energy thus states that

$$\delta Q = \delta U + \delta W \quad (1.7)$$

Experiments are concerned only with differential increments which are not perfect differentials of the functions Q and W of the coordinates or variables specifying the system. A perfect differential of a function f of the variables x_i is given by $df = \sum \frac{\partial f}{\partial x_i} dx_i$ and $\int_1^2 df = f_2 - f_1$. That is, a function f exists of which df is the differential. The condition that an arbitrary differential expressed as a function of two variables, say x and y , such as $\delta L = M(x,y) dx + N(x,y) dy$ shall be a perfect differential, i.e., that a function L shall exist, is $\delta L = dL$, or

$$M(x,y) = \frac{\partial L}{\partial x}, \quad N(x,y) = \frac{\partial L}{\partial y}$$

Thus the test is the satisfaction of the condition

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

It may be shown that in the case of two variables an integrating factor always exists whereby the differential equation may be converted into one between perfect differentials, and a significant function of the physical variables is thus defined. In the example above, if λ is taken to be the integrating factor, $\lambda \delta L$ is a perfect differential if $\partial(\lambda M)/\partial y = \partial(\lambda N)/\partial x$. This equation determines λ .

These considerations may be used to define useful thermodynamic functions, analogous to the potential-energy functions of mechanics, from Eq. (1.7). Choosing the integrating factor, for reasons that will later become evident, as $1/T$ and for purposes of specificity considering a gas doing work $\delta W = p dV$ against an external pressure,

$$\frac{\delta Q}{T} = \frac{dU}{T} + \frac{p dV}{T} = dS \quad (1.8)$$

δU is written dU since U is a function of the variables defining the system and its variation is a perfect differential. Equation (1.8) is the definition of the perfect differential dS of the postulated function S which is known as the *entropy*. The reason for the particular choice of the integrating factor is that extensive experience with gases has shown that in their behavior they approximate a hypothetical substance known as a perfect gas which may be taken by definition to have the following two properties:

$$pV = RT \quad \text{and} \quad dU = C_V dT$$

where C_V (the specific heat at constant volume) and R (the gas constant per mole) are constants. The T that occurs here is the absolute or Kelvin temperature. Using these equations to eliminate p and dU , it is evident that dS is a perfect differential if the integrating factor is the reciprocal of the absolute temperature:

$$dS = C_V \frac{dT}{T} + R \frac{dV}{V}$$

or

$$S = C_V \ln T + R \ln V + \text{constant}$$

The definition of T is that it is the integrating factor of Eq. (1.7), and the illustration above is simply for the purpose of identifying it with the parameter that occurs in the two expressions which characterize the perfect gas.

A physical process is said to take place isentropically if the conditions of thermal insulation are such that there is no heat exchange so that δQ and dS are zero. Similarly, an isovoluminal change is one for which $dV = 0$. From Eq. (1.8), dU may be written

$$dU = T dS - p dV$$

and for an isovoluminal isentropic process $dU = 0$, or $U(S, V) = \text{constant}$. Thus U is an expression for the energy in terms of the entropy and volume which plays the same role in thermodynamic problems as does H , the mechanical energy, in mechanical problems. The analogues of the mechanical equations of motion are $(\partial U / \partial S)_V = T$ and $(\partial U / \partial V)_S = -p$. Other functions analogous to U can be constructed suitable for use in circumstances in which any two of the four variables S , T , V , and p are maintained constant. These functions then are useful in the determination of thermodynamic equilibria because of their constancy under the imposed conditions.

Internal energy U : $dU = T dS - p dV$, $U(S, V) = \text{constant}$

Enthalpy $x = U + pV$: $d_x = T dS + V dp$, $x(S, p) = \text{constant}$

Helmholtz free energy $\Psi = U - TS$: $d\Psi = -S dT - p dV$,

$\Psi(T, V) = \text{constant}$

Gibbs free energy $\Phi = U - TS + pV$: $d\Phi = -S dT + V dp$,

$\Phi(T, p) = \text{constant}$

The above synopsis of the basis of thermodynamics indicates the concepts that it employs and the analogy between the energy method of handling mechanical problems and the corresponding method for thermodynamic problems. As an instance, consider the relationship between the temperature and pressure required for equilibrium of a vapor (desig-

nated by the subscript 1) in contact with a liquid of the same substance (subscript 2). The equilibrium condition for constant temperature and pressure is that $d\Phi = 0$ for a transfer of some of the material between the liquid and vapor phases.

$$d\Phi = -(S_1 - S_2) dT + (V_1 - V_2) dp = 0$$

$$\left(\frac{\partial p}{\partial T}\right)_\Phi = \frac{S_1 - S_2}{V_1 - V_2}$$

If the heat of condensation per mole is Q_T , $Q_T/T = S_1 - S_2$ and

$$\left(\frac{\partial p}{\partial T}\right)_\Phi = \frac{Q_T}{T(V_1 - V_2)}$$

This equation (*Clapeyron's*) describes the behavior of vapor pressure in terms of temperature when the heat of change of state Q and the volumes per mole of the states V_1 and V_2 are known.

1.6. Relativistic Mechanics. In the formulation of Newtonian mechanics the temporal variable occupies a unique position as the independent variable in terms of which the position of a body is described. It is fundamentally unrelated to the changes in coordinate systems which it may be convenient to introduce for the description of mechanical problems. If a body is moving along the z axis with a velocity v , the description of phenomena at rest with respect to the body would be in terms of the primed axes of Fig. 1.6 instead of the unprimed axes

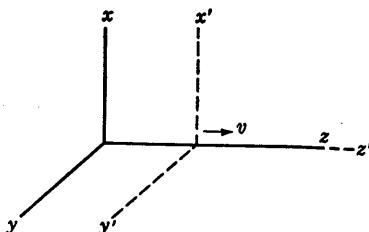


FIG. 1.6. Relative motion of cartesian coordinate systems.

which are at rest with respect to another observer. The relationships between the coordinate systems are

$$\begin{aligned} x &= x', & z &= z' + vt \\ y &= y', & t &= t' \end{aligned}$$

The velocity of a body or disturbance as measured in one coordinate system is simply the vectorial sum of the velocity in a second coordinate system and the velocity of the second system with respect to the first. The type of system in which such a description is in accord with experience is known as a *Galilean system*, and such a concept proves adequate for most of the common phenomena of classical mechanics.

However, it is found by suitably refined observation that this approximation provides an inadequate description of phenomena. In particular, the Michelson-Morley¹ type of experiment demonstrates that when a

¹ A. A. Michelson and E. R. Morley, *Am. J. Sci.*, **34**, 333 (1887).

measurement is made of the velocity of light or any other electromagnetic disturbance in free space a constant value is obtained regardless of the motion of the observer. In this experiment the interference pattern of two light beams which traverse paths at right angles to one another is observed while the point on the earth's surface at which the observation is being made changes its velocity with the motion of the earth and solar system. Since the pattern is found to display no change within the limits of error of the measurement (the precision of the experiment is very great in the Kennedy-Thorndyke modification¹), it may be concluded as an experimental fact that the velocity of light is a constant in our universe independent of the velocity of the observer. If the letter c is written for the velocity of light, the condition to be satisfied becomes

$$c^2 = \frac{x^2 + y^2 + z^2}{t^2} = \frac{x'^2 + y'^2 + z'^2}{t'^2} \quad (1.9)$$

Thus a relation exists between the spatial and temporal variables used by any observer which constitutes an entirely new experimental fact requiring modification of the simple Galilean equations of transformation.

The rectilinear propagation of light as observed experimentally requires that the relationship between the primed and unprimed variables be a linear one. For convenience in nomenclature let $x = x_1$, $y = x_2$, $z = x_3$, and $ict = x_4$ (where $i = \sqrt{-1}$) and similarly for the primed variables (see Fig. 1.7). Then the general linear transformation for the four variables ($i \dagger$ and j from 1 to 4) may be written

$$x_i = \sum_j a_{ij} x'_j \quad (1.10)$$

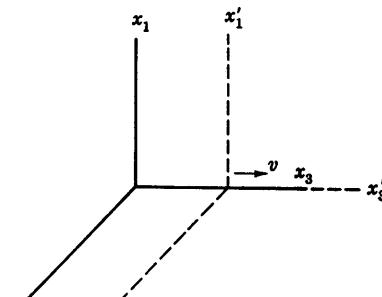


FIG. 1.7. Coordinate nomenclature.

Equations (1.9) state that the square of the radius vector of four-dimensional space, $x^2 + y^2 + z^2 - c^2 t^2$ or $x'^2 + y'^2 + z'^2 - c^2 t'^2$, is invariant under a transformation of coordinates, just as $x^2 + y^2 + z^2$ is a square of an invariant length in ordinary three-dimensional statics. The primed and unprimed expressions need not actually be equal to one another so far as the evidence hitherto cited is concerned. The primed and unprimed

¹ R. J. Kennedy and E. M. Thorndyke, *Phys. Rev.*, **42**, 400 (1932).

† The use of i for a running index is not to be confused with the i as the square root of minus one.

systems could differ by a scale factor. However, an experiment by Ives and Stilwell¹ on the Doppler effect of a moving light source shows that they are actually equal, and hence

$$\sum_i x_i^2 = \sum_i x_i'^2 \quad (1.11)$$

Therefore, squaring and adding Eqs. (1.10), it is seen that

$$\sum_i a_{ij}a_{ik} = \delta_{jk} \quad (1.12)$$

where $\delta_{jk} = 1$ if $j = k$, and $\delta_{jk} = 0$ if $j \neq k$. Equation (1.11) shows that there is no general contraction or dilatation with a change to another set of moving axes, and it can be assumed that two sets of axes perpendicular to the direction of motion are unchanged, i.e.,

$$\begin{aligned} x_1 &= x'_1, & x_2 &= x'_2 \\ x_3 &= a_{33}x'_3 + a_{34}x'_4, & x_4 &= a_{43}x'_3 + a_{44}x'_4 \end{aligned}$$

where x_3 is chosen as the axis along which the origin of the primed axis is assumed to move with the velocity v .

Applying Eqs. (1.12),

$$a_{33}^2 + a_{43}^2 = a_{34}^2 + a_{44}^2 = 1 \quad \text{and} \quad a_{33}a_{34} + a_{43}a_{44} = 0$$

These three equations among the four coefficients may be used to reduce them to a single parameter, say, $a_{33} = \cos \phi$. Then

$$\begin{aligned} x_1 &= x'_1, & x_2 &= x'_2 \\ x_3 &= \cos \phi x'_3 + \sin \phi x'_4, & x_4 &= -\sin \phi x'_3 + \cos \phi x'_4 \end{aligned}$$

Applying then the condition that the origin of the primed system ($x'_3 = 0$) moves with a velocity v in the unprimed system,

$$\frac{x_3}{x_4} = \tan \phi \quad \text{or} \quad \phi = \tan^{-1} \frac{z}{ict} = \tan^{-1} \frac{v}{ic}$$

Hence the transformation equations, which are known as the *Lorentz transformation*, are, in terms of the x , y , and z nomenclature,

$$\begin{aligned} x &= x', & y &= y' \\ z &= \frac{z' + vt'}{\sqrt{1 - v^2/c^2}}, & t &= \frac{t' + vz'/c^2}{\sqrt{1 - v^2/c^2}} \end{aligned} \quad (1.13)$$

These equations are seen to reduce approximately to the Galilean equations for small v/c and hence to be in accord with familiar observation.

¹ H. E. Ives and G. R. Stilwell, *J. Opt. Soc. Amer.*, **28**, 215 (1938); **31**, 369 (1941).

By forming the appropriate derivatives of Eqs. (1.13) the equations for transformation of velocities and accelerations can be obtained, thus providing a consistent kinematics. Our concern is particularly with the effect of these ideas on dynamics. As a basic postulate in agreement with experience it is assumed that the laws of dynamics are of the familiar form for any observer. In any coordinate system the total momentum of an isolated system is an invariant quantity, that is, $\mathbf{p} = m\mathbf{v}$ is conserved, though of course the measured value of this quantity will depend upon the relative motion of the observer and the mass. In particular, an observer at rest with respect to a mass finds by experiments with balances or springs that it is invariant with time, i.e., it does not change in comparison with other masses in the same reference system. The time measured by a clock at rest with respect to a mass is known as the *proper time*.

Consider a situation in which an elastic ball of mass m_a , measured to be m_0 at rest in the unprimed system, is thrown along the y direction with a velocity \mathbf{u} so as to make a collision at $y = 0$, $x = 0$, and $z = 0$ at a time $t = 0$ with a similar ball m_b measured to be m_0 at rest in the primed system (which is moving with velocity \mathbf{v} in the z direction relative to the unprimed system). m_b has similarly been thrown along the y' axis with a velocity $-\mathbf{u}$ measured in the primed system. After collision, these must rebound with reversed velocities equal to initial velocities in consideration of the symmetry of the arrangement. The total momentum of the balls must be conserved when measured in either system. Consider the measurements in the unprimed system. The y component of velocity of mass m_a is initially $w_{ay} = u$ and finally $w_{ay} = -u$. The y component of the velocity of mass m_b is initially $w_{by} = -u(1 - v^2/c^2)^{\frac{1}{2}}$ and finally $w_{by} = u(1 - v^2/c^2)^{\frac{1}{2}}$. The z components of these velocities are $w_{az} = 0$ and $w_{bz} = v$. Conservation of momentum in the y direction gives

$$2m_a u = 2m_b \left(1 - \frac{v^2}{c^2} \right)^{\frac{1}{2}}$$

If u becomes very small, m_a will approach m_0 , and m_b , written simply m , is seen to be necessarily a function of v .

$$m(v) = m_0 \left(1 - \frac{v^2}{c^2} \right)^{-\frac{1}{2}}$$

A similar conclusion is reached if the collision is considered in the primed coordinate system. Thus the apparent mass depends upon the velocity between the mass and the observer. The greater the velocity, the larger the mass—which tends to infinity as the velocity approaches c . This relativistic increase in mass is confirmed by precise measurements of

the specific masses of electrons at high velocities and also by all the recent experience with high-energy-particle accelerators.

Since it is postulated that the form of the equations of motion,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}$$

is invariant for all observers, another conservation law is found to exist which is an extension of the Newtonian conservation of energy. Assuming for purposes of illustration the existence of a potential-energy function $V(z)$ the equation of motion is

$$\frac{d}{dt} \left(\frac{m_0 v}{\sqrt{1 - v^2/c^2}} \right) = - \frac{dV}{dz} \quad (1.14)$$

Multiplying both sides by $v = dz/dt$, the velocity with which m_0 is moving in the z direction, and performing the differentiation allowing for the variation of v ,

$$\frac{m_0 v \, dv/dt}{(1 - v^2/c^2)^{3/2}} = \frac{d}{dt} \left(\frac{m_0 c^2}{\sqrt{1 - v^2/c^2}} \right) = - \frac{dV}{dt}$$

or

$$\frac{d}{dt} (mc^2 + V) = 0 \quad (1.15)$$

Thus the quantity $mc^2 + V$ is a constant of the motion, and hence, if mc^2 is taken to be the energy, Eq. (1.15) is an extension of the statement of the conservation of energy concordant with the special theory of relativity. It is also in accord with Newtonian mechanics if v/c is small enough so that second-order terms can be neglected since

$$mc^2 = m_0 c^2 + \frac{1}{2} m_0 v^2 + \dots$$

mc^2 being merely an additive constant. Thus the laws of conservation of momentum, $\mathbf{p} = m\mathbf{v}$, and of energy, $E = mc^2$, are adequate descriptions of all mechanical phenomena in accord with experimental observation under the conditions that have been assumed.

The quantities \mathbf{p} and E are clearly related as follows:

$$\left(\frac{E}{c} \right)^2 - p^2 = (m_0 c)^2 = \text{constant} \quad (1.16)$$

In the form

$$p_x^2 + p_y^2 + p_z^2 - \left(\frac{E}{c} \right)^2 = p'_x^2 + p'_y^2 + p'_z^2 - \left(\frac{E'}{c} \right)^2$$

the equation is seen to be of the same form as Eqs. (1.9), and hence by a

precisely similar argument as led to the coordinate transformation [Eqs. (1.13)] it is seen that

$$\begin{aligned} p_x &= p'_x, & p_y &= p'_y \\ p_z &= \frac{p'_z + E'v/c^2}{\sqrt{1 - v^2/c^2}}, & E &= \frac{E' + vp'_z}{\sqrt{1 - v^2/c^2}} \end{aligned}$$

These equations describe the transformation of momentum and energy between moving coordinate systems.

1.7. Electricity and Electromagnetism. The concepts of electricity, electrification, and the conservation of electric charge are bound up with the experimental observations of electrostatics. The quantity called electric charge is analogous in some respects to mass, in that charges are observed to exert forces on one another as do masses. Charges, however, differ from masses in that charges can be either positive or negative in the sense of Eq. (1.2), i.e., charges may either attract or repel one another. The law of force between charges, Coulomb's law, is of the same form as that between masses, and in terms of the usual unit of charge, the coulomb, the law is

$$\mathbf{F}_{q_1} = -\frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}^2} \mathbf{r}_{12}$$

where $\epsilon_0 = 8.85 \times 10^{-12}$ coulomb² joule⁻¹ m⁻¹ and \mathbf{r}_{12} is a unit vector from q_1 to q_2 . The auxiliary concepts of fields and potentials are particularly useful in electricity to describe the nature of the motion of a charged test body in terms of the properties of neighboring space. The *electric field* is the force per unit charge, assuming negligible disturbance of conditions by the test charge, and the *electric potential* is defined as the scalar quantity whose negative gradient is the electric field.

It was found by Ampère and Faraday that other forces appear if the charges which are present are moving relative to an observer. To describe the force on a charge in the presence of moving charges, a new vector concept must be introduced which is generally known as *magnetic induction*.¹ It is not the negative gradient of a scalar quantity but is defined as the vector which, when multiplied vectorially by the velocity of a moving test charge, yields the force per unit charge that is observed. Thus the total force on a test charge e can be written

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (1.17)$$

¹ This nomenclature, imposed by electrical-unit committees, is somewhat misleading physically, particularly in association with the commonly used terms induced magnetism, induced emf, etc. The magnetic induction \mathbf{B} is the assumed vector field associated with the magnetic force on a moving test charge.

where \mathbf{E} and \mathbf{B} are the electric field and magnetic induction, respectively. The equations which state the relationships between these vectors and the motions of the charges throughout space that give rise to them are known as Maxwell's equations. These represent the laws of Coulomb, Ampère, Faraday, and the conservation of charge. In units of meters, seconds, and coulombs these equations are

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= - \frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} &= \frac{1}{c^2} \left(\frac{\rho \mathbf{v}}{\epsilon_0} + \frac{\partial \mathbf{E}}{\partial t} \right)\end{aligned}\quad (1.18)$$

where $\nabla = \mathbf{i} \partial/\partial x + \mathbf{j} \partial/\partial y + \mathbf{k} \partial/\partial z$ in rectangular coordinates, ρ is the charge density, and c is a constant having the dimensions of velocity.

Many of the important applications of electrical concepts to atomic theory involve only static or quasi-static conditions. Thus the potential energy of a static configuration of charges, which is obtained directly by performing the integration $\int \mathbf{F} \cdot d\mathbf{r}$ using the Coulomb law, is

$$U_e = \frac{1}{8\pi\epsilon_0} \sum_{i,j} \frac{q_i q_j}{r_{ij}} = \frac{1}{2} \sum_i q_i \varphi_i \quad i \neq j$$

where the electric potential φ_i at charge i due to all other charges is given by $\varphi_i = \frac{1}{4\pi\epsilon_0} \sum_j \frac{q_j}{r_{ij}}$ (see Fig. 1.8).

An important quasi-static concept is that of *magnetic moment*, which embodies the electromagnetic properties associated with a loop of electric current. The elementary magnetic dipole moment may be thought of as a small circular current of constant magnitude and dimensions or as an electric charge e circulating at a constant angular velocity ω in a small circular orbit of constant area a . The general definition of the magnetic moment of a charge e at a position \mathbf{r}' moving with a velocity \mathbf{v} is

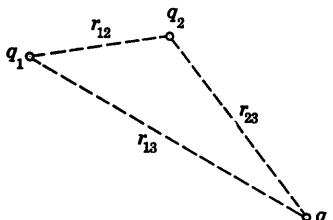


FIG. 1.8. Electrostatic energy of point charges.

$$\mathbf{u} = \frac{1}{2} e \mathbf{r}' \times \mathbf{v}$$

From Fig. 1.9 it is seen that $\frac{1}{2} \mathbf{r}' \times \mathbf{v}$ is the rate at which the vector \mathbf{r}'

describes an area. For a circular orbit, which is most useful in atomic theory, described at the angular rate ω ,

$$\mu = \frac{ea}{2\pi} \omega$$

The concept of magnetic moment, as in Fig. 1.9, is not limited to such a specific description though it is useful in calculating the magnetic induction due to μ from Eqs. (1.18) and in calculating the forces and torques affecting μ in a magnetic induction \mathbf{B} . Indeed, the concept of magnetic moment is extended in certain instances to describe phenomena that cannot be thought of as due to a charge in rotational motion, and the moment associated with it is the basic datum entering the theory. The magnetic induction produced by a magnetic dipole moment μ at a distant point \mathbf{r} may be shown from Eqs. (1.18) to be

$$\mathbf{B} = \frac{-1}{4\pi\epsilon_0 c^2} \nabla \times \left(\mu \times \nabla \frac{1}{r} \right)$$

The potential energy of a magnetic moment μ in a region of magnetic induction \mathbf{B} may be taken as

$$U_m = -\mu \cdot \mathbf{B}$$

in the sense that the negative gradient of U_m yields the force and torque on μ due only to \mathbf{B} .

Most practical problems in the field of electricity that involve electric currents are conveniently handled in terms of a quantity known as the *magnetic vector potential*, \mathbf{A} , which is defined by

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (1.19)$$

From the third of Eqs. (1.18), $\nabla \times (\mathbf{E} + \partial \mathbf{A} / \partial t) = 0$ or, since

$$\nabla \times \nabla \varphi = 0$$

where φ is any scalar, $-\mathbf{E} = \partial \mathbf{A} / \partial t + \nabla \varphi$; hence φ corresponds to the electrostatic potential since, when conditions are constant in time, $\mathbf{E} = -(\nabla \varphi)$. The second of Eqs. (1.18) is identically satisfied by Eq. (1.19). If $\nabla \cdot \mathbf{A}$, which is as yet undefined, is taken to be equal to $-(1/c^2) \partial \varphi / \partial t$, the first and fourth equations of Eqs. (1.18) become

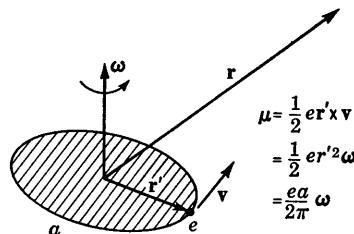


FIG. 1.9. Magnetic-moment concepts.

$$\begin{aligned} -\nabla^2\varphi + \frac{1}{c^2} \frac{\partial^2\varphi}{\partial t^2} &= \frac{\rho}{\epsilon_0} \\ -\nabla^2\mathbf{A} + \frac{1}{c^2} \frac{\partial^2\mathbf{A}}{\partial t^2} &= \frac{1}{c^2} \frac{\rho\mathbf{v}}{\epsilon_0} \end{aligned} \quad (1.20)$$

In free space the net charge density vanishes and Eqs. (1.20) are seen to be the differential equations of wave motion representing a disturbance propagated with the velocity c . Such a disturbance is known as an electromagnetic wave. In free space the vectors \mathbf{E} and \mathbf{B} also satisfy equations similar to Eq. (1.20) with zero on the right.

A solution of Laplace's equation for φ , which is the first of Eqs. (1.20) when there is no time variation, is known from the Coulomb law to be

$$\frac{1}{4\pi\epsilon_0} \sum_i \frac{q_i}{r_i} \text{ or, in terms of a charge density } \rho \text{ and the corresponding integral}$$

over the volume occupied by it, $\frac{1}{4\pi\epsilon_0} \int \frac{\rho}{r} d\tau$ where $d\tau$ is the element of volume. The term involving the time in Eqs. (1.20) results in delaying the effect of the presence of a charge at a distant point by the time taken by an electromagnetic wave to travel between the charge and the point of observation. Thus it can be shown that a solution of the first of Eqs. (1.20) can be written

$$\varphi = \frac{1}{4\pi\epsilon_0} \int \frac{((\rho))}{r} d\tau$$

where the double parentheses indicate that in forming the integral ρ is to be taken as that value of the charge density existing at a time earlier by r/c than the time in terms of which φ is expressed, that is, $((\rho)) = \rho(t - r/c)$. By analogy a solution of the second of Eqs. (1.20) is

$$\mathbf{A} = \frac{1}{4\pi\epsilon_0 c^2} \int \frac{((\rho\mathbf{v}))}{r} d\tau$$

In most problems of interest in atomic physics the linear dimensions of the volume occupied by the charge distribution giving rise to φ and \mathbf{A} are small compared to the distance to the point at which φ and \mathbf{A} produce the effects to be described. Under these circumstances a convenient type of series expansion in terms of a charge distribution approximated by dipoles, quadrupoles, and higher multipoles can be used to advantage (see Sec. 2.6).

The emission, absorption, and scattering of electromagnetic radiation by matter are all very important processes in which atoms participate. These processes are not strictly amenable to a classical treatment. Nevertheless, there are many qualitative points of similarity between the

interaction of radiation and charges on an electronics or engineering scale and the interaction of high-frequency radiation and atomic systems. Therefore, a brief review of the salient points of classical dipole-radiation theory is very useful as background for the appreciation of atomic processes in which radiation is involved. The emission and absorption of radiation by atoms can be associated with the alteration of a charge configuration from an initial to a final form, and the features such as intensity and polarization are describable in terms of these charge configurations just as in classical processes. Again, the elastic and inelastic scattering processes can be described qualitatively in classical terms, and even the quantitative results are not greatly in error at relatively great wavelengths. In consequence of the importance of classical radiation theory as background in describing atomic processes, a brief account of the elementary processes involved in dipole radiation is given below and in the following section.

Assume the presence of a dipole (Fig. 1.10) described by $\mathbf{P} = el$, where l is the vector separation of charges $+e$ and $-e$, and assume that \mathbf{P} is a periodic function of the time given in complex terms by $\mathbf{P} = \mathbf{P}_0 e^{i\omega t}$, where $\mathbf{P}_0 = el_0$. \mathbf{P} is known as the *dipole moment* and P_0 as its amplitude. Then if \mathbf{r} is measured from the mid-point of these charges and l_0 is small enough so that $(l_0/r)^2$ can be neglected,

$$\begin{aligned}\varphi &= \frac{e}{4\pi\epsilon_0} \left(\frac{e^{i(\omega t - kr + \frac{1}{2}kl_0 \cos \theta)} - e^{i(\omega t - kr - \frac{1}{2}kl_0 \cos \theta)}}{r} \right) \\ &= \frac{1}{4\pi\epsilon_0} 2ie \sin \left(\frac{kl_0 \cos \theta}{2} \right) \frac{e^{i(\omega t - kr)}}{r}\end{aligned}$$

where k is written for ω/c . If kl_0 is also small, that is, l_0 is much less than the wavelength of the radiation, the sine can be replaced by the argument and

$$\varphi = \frac{ik}{4\pi\epsilon_0} \frac{e^{i(\omega t - kr)}}{r} \mathbf{P}_0 \cdot \mathbf{r}_1$$

Similarly, since $\rho v d\tau = e\dot{l}/2 - (-e\dot{l}/2) = i\omega\mathbf{P}$, where the dot is used to indicate the time derivative, the expression for \mathbf{A} may be written

$$\mathbf{A} = \frac{i\omega}{4\pi\epsilon_0 c^2} \frac{e^{i(\omega t - kr)}}{r} \mathbf{P}_0$$

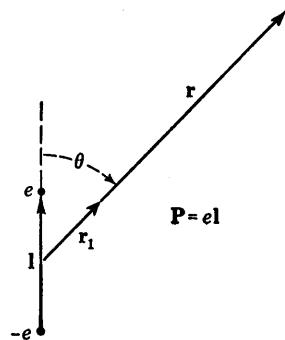


FIG. 1.10. Electric dipole.

Thus \mathbf{A} is seen to be a vector which is parallel to the dipole moment (Fig. 1.11), which varies with the time as $e^{i\omega t}$, and which varies with r as $\frac{e^{-ikr}}{r}$. The expressions for φ and \mathbf{A} above correspond to what is called a spherical wave diverging from the dipole.

The corresponding magnetic induction due to the dipole is given by

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{-i\omega}{4\pi\epsilon_0 c^2} \left(ik + \frac{1}{r} \right) \frac{e^{i(\omega t - kr)}}{r} \mathbf{r}_1 \times \mathbf{P}_0$$

where \mathbf{r}_1 is a unit vector in the direction \mathbf{r} . If, further, the point of observation is many wavelengths away, that is, $1/r \ll k$, the second term in the parentheses can be neglected and

$$\mathbf{B} = \frac{\omega^2}{4\pi\epsilon_0 c^3} \frac{e^{i(\omega t - kr)}}{r} \mathbf{r}_1 \times \mathbf{P}_0 \quad (1.21)$$

The associated electric field can be found from φ or more simply from the fourth of Eqs. (1.18) (with $\rho = 0$). Since the time variation is assumed to be given by $e^{i\omega t}$, $(1/c^2) \partial \mathbf{E} / \partial t$ is $i\omega \mathbf{E}/c^2$, and $\nabla \times \mathbf{B}$ is by an analogous argument equal to $-ik\mathbf{r}_1 \times \mathbf{B}$. Thus

$$\mathbf{E} = -c\mathbf{r}_1 \times \mathbf{B} \quad (1.22)$$

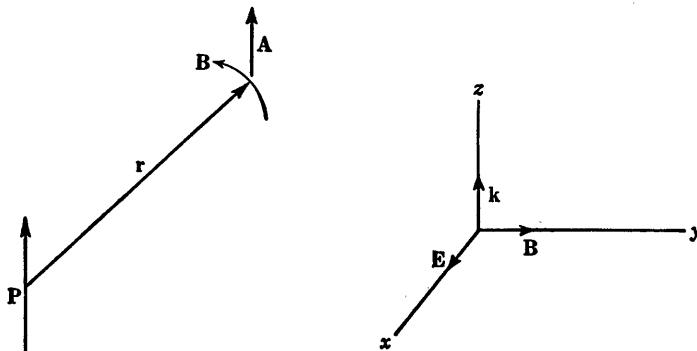


FIG. 1.11. Magnetic vectors associated with an oscillating electric dipole.

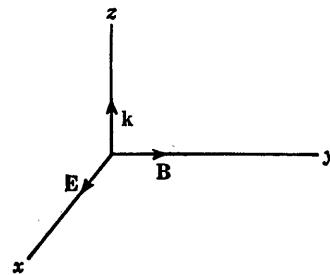


FIG. 1.12. Disposition of wave vectors.

1.8. Electromagnetic Waves and Oscillators. The electromagnetic wave represented by \mathbf{B} and \mathbf{E} is propagated away from the dipole in the direction \mathbf{r}_1 which may alternatively be designated by the vector \mathbf{k} (Fig. 1.12). It carries energy with it as may be seen by considering the force it exerts on a surface containing charges free to move under the

influence of the force. From the identity

$$\frac{\partial(\mathbf{E} \times \mathbf{B})}{\partial t} = \frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B} + \mathbf{E} \times \frac{\partial \mathbf{B}}{\partial t}$$

and the third and fourth of Eqs. (1.18), we have

$$\mathbf{v} \times \mathbf{B} = -\frac{\epsilon_0}{\rho} \left[c^2 \mathbf{B} \times (\nabla \times \mathbf{B}) + \mathbf{E} \times (\nabla \times \mathbf{E}) + \frac{\partial(\mathbf{E} \times \mathbf{B})}{\partial t} \right]$$

Using Eq. (1.17),

$$\mathbf{F}_{\text{vol}} = \left[\rho \mathbf{E} - \epsilon_0 c^2 \mathbf{B} \times (\nabla \times \mathbf{B}) - \epsilon_0 \mathbf{E} \times (\nabla \times \mathbf{E}) - \epsilon_0 \frac{\partial(\mathbf{E} \times \mathbf{B})}{\partial t} \right]$$

If there is no net charge ($\rho = 0$) and the vectors \mathbf{B} and \mathbf{E} are given by Eqs. (1.21) and (1.22), the identity $\mathbf{B} \times (\nabla \times \mathbf{B}) = \frac{1}{2} \nabla(B^2) - (\mathbf{B} \cdot \nabla) \mathbf{B}$ and the similar one for \mathbf{E} may be used to write the force:

$$\mathbf{F}_{\text{vol}} = - \left[\frac{1}{2} \epsilon_0 c^2 \nabla(B^2) + \frac{1}{2} \epsilon_0 \nabla(E^2) + \epsilon_0 \frac{\partial(\mathbf{E} \times \mathbf{B})}{\partial t} \right]$$

If the average value is taken over a considerable time, the last term vanishes and

$$\bar{\mathbf{F}}_{\text{vol}} = -\nabla \left(\frac{\epsilon_0 c^2}{2} B^2 + \frac{\epsilon_0}{2} E^2 \right) \quad (1.23)$$

The quantity in the parentheses plays the role of an energy density since the negative gradient of it gives the force per unit area exerted by the wave. Finally, considering a volume of free space out of which an electromagnetic wave is being propagated, and forming the rate of loss of energy,

$$\begin{aligned} -\frac{\partial}{\partial t} \left(\frac{\epsilon_0 c^2}{2} B^2 + \frac{\epsilon_0}{2} E^2 \right) &= - \left(\epsilon_0 c^2 \mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} + \epsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} \right) \\ &= \epsilon_0 c^2 (\mathbf{B} \cdot \nabla \times \mathbf{E} - \mathbf{E} \cdot \nabla \times \mathbf{B}) \\ &= \epsilon_0 c^2 \nabla \cdot (\mathbf{E} \times \mathbf{B}) \end{aligned}$$

Integrating this over a volume, the left-hand side gives the rate of decrease of energy in the volume and the right-hand side, by Gauss's theorem, is the integral of the normal component of the vector $\epsilon_0 c^2 \mathbf{E} \times \mathbf{B}$ over the bounding area. Thus the vector

$$\mathbf{N} = \epsilon_0 c^2 \mathbf{E} \times \mathbf{B}$$

which is known as *Poynting's vector*, may be taken to correspond to a rate of flow of energy per unit area out through the bounding surface.

Taking \mathbf{E} and \mathbf{B} from Eqs. (1.21) and (1.22) and choosing the cosine for the real part of the exponential, the energy per unit area at colatitude θ flowing through a sphere of radius r surrounding the oscillating dipole (Fig. 1.13) is

$$\begin{aligned}\mathbf{N}(\theta, r) &= \epsilon_0 c^3 \mathbf{B} \times (\mathbf{r}_1 \times \mathbf{B}) = \epsilon_0 c^3 B^2 \mathbf{r}_1 \\ &= \frac{\omega^4}{16\pi^2 \epsilon_0 c^3} \frac{\cos^2(\omega t - kr)}{r^2} P_0^2 \sin^2 \theta \mathbf{r}_1\end{aligned}$$

Thus the outflow of energy is greatest in the equatorial plane, where $\theta = \pi/2$, and least toward the poles, where $\theta = 0$ or π . The rate of loss of energy over the entire sphere is obtained by multiplying N by

$$2\pi r^2 \sin \theta d\theta$$

and integrating from $\theta = 0$ to $\theta = \pi$.

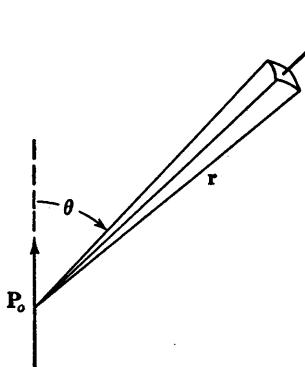


FIG. 1.13. Radiation from a dipole.

$$N = \frac{\omega^4}{8\pi\epsilon_0 c^3} \cos^2(\omega t - kr) P_0^2 \int_0^\pi \sin^3 \theta d\theta$$

As the integral equals $\frac{4}{3}$,

$$N = \frac{\omega^4 P_0^2}{6\pi\epsilon_0 c^3} \cos^2(\omega t - kr) \quad (1.24)$$

The time average of $\cos^2(\omega t - kr)$ is $\frac{1}{2}$, so that the average rate of radiation of energy by the dipole out in all directions is

$$N = \frac{\omega^4 P_0^2}{12\pi\epsilon_0 c^3} = 1.11 \times 10^{-16} P_0^2 \omega^4 \quad \text{watts}$$

where P_0 is in coulomb-meters and ω in radians per second.

The rate of radiation of energy given by Eq. (1.24) may be interpreted in terms of a retarding force against which any energy source maintaining the oscillatory motion would have to do work. If the dipole moment at the origin, $r = 0$, is taken as being in the z direction, \ddot{P} may be written

$$\ddot{P} = e\ddot{z} = -\omega^2 P_0 \cos \omega t$$

and N may be expressed in terms of z as

$$N = \frac{1}{6\pi\epsilon_0 c^3} (e\ddot{z})^2$$

This rate of radiation of energy may be equated to the rate at which an

energy source does work against a retarding force F_r , and the result is integrated over a complete period of oscillation.

$$\int_{\tau} F_r \dot{z} dt = \frac{e^2}{6\pi\epsilon_0 c^3} \int_{\tau} \ddot{z}^2 dt$$

The right-hand side may be integrated by parts to yield

$$\int_{\tau} F_r \dot{z} dt = \frac{e^2}{6\pi\epsilon_0 c^3} \left(|\dot{z}\ddot{z}|_{\tau} - \int_{\tau} \ddot{z}\dot{z} dt \right)$$

If the energy source maintains the amplitude of oscillation, $\dot{z}\ddot{z}$ has the same value at the beginning and end of a period and in consequence the integrands can be equated, yielding

$$F_r = - \frac{e^2}{6\pi\epsilon_0 c^3} \ddot{z}$$

In the case of simple harmonic motion which is here being considered, $z = z_0 e^{i\omega_0 t}$ and \ddot{z} may be written as $-\omega_0^2 z$. In these terms the radiation force is closely analogous to that of ordinary viscous damping. The mechanical equation of motion of a free oscillator of mass m and of natural frequency ω_0 would be written

$$m\ddot{z} + m\omega_0^2 z - \frac{e^2}{6\pi\epsilon_0 c^3} \ddot{z} = 0$$

or

$$\ddot{z} + \gamma \dot{z} + \omega_0^2 z = 0 \quad (1.25)$$

where

$$\gamma = \frac{2}{3} \frac{\omega_0^2}{c} r_0 \quad \text{and} \quad r_0 = \frac{e^2}{4\pi\epsilon_0 m c^2} \quad (1.26)$$

Thus the damping force is proportional to ω_0^2 , and the constant r_0 is evidently equal to the separation of two charges e at which the electrostatic energy is equal to the relativistic energy equivalent to the mass m . If e and m refer to an electron, r_0 is known as the *classical electron radius* and it is equal to 2.82×10^{-15} m. As is brought out in the subsequent discussion γ is equal to the width of the response curve at half maximum and hence is an inverse measure of the sharpness of resonance of the oscillator.

The preceding argument dealing with an oscillating dipole can be extended immediately to the case of a mass m of charge e rotating in a circular orbit of constant radius about a stationary charge of equal but opposite sign as shown in Fig. 1.14. The motion is evidently equivalent to that of two equal dipoles with their axes at right angles to one another

and differing by $\pi/2$ in temporal phase. Thus

$$P_x = P_{x0} \cos \omega_0 t, \quad P_y = P_{y0} \sin \omega_0 t,$$

or

$$P = P_x + iP_y = P_0 e^{i\omega_0 t} \quad (1.27)$$

The magnitudes and phase relations of the field vectors at a distant point can be immediately computed from Eqs. (1.21) and (1.22). In general

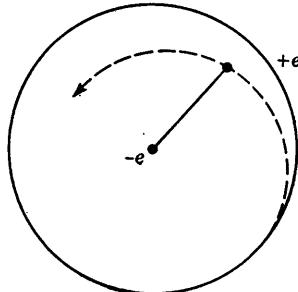


FIG. 1.14. Radiation from a circulating charge.

the radiation is elliptically polarized; it is circularly polarized when viewed along the z axis and linearly polarized when observed in the xy plane. The point of particular interest is that such an oscillator possesses the mechanical angular momentum $mr^2\omega$ in addition to the kinetic energy $\frac{1}{2}mr^2\omega^2$. If it radiates energy by slowing down slightly, it loses both energy and angular momentum which by the basic conservation hypotheses must be carried away by the radiation. The rate of energy loss is $mr^2\omega d\omega/dt$, and the rate of angular momentum loss is $mr^2 d\omega/dt$, so such an oscillator radiates angular momentum at $1/\omega$ the rate at which it radiates energy.

The scattering and absorption of an electromagnetic wave by a free charge or a dipole are processes of interest in atomic physics. The free-charge case can be considered as a special one in which the restoring force of the oscillator and its internal damping force vanish. The equation of free motion of a damped oscillator is given by Eq. (1.25). To a first approximation the force exerted by an electromagnetic wave is eE , or (with the understanding that the real part of the exponential is to be taken) the driving force per unit mass which replaces zero on the right of Eq. (1.25) is $(eE_0/m)e^{i\omega t}$, and

$$\ddot{z} + \gamma \dot{z} + \omega_0^2 z = \frac{eE_0}{m} e^{i\omega t}$$

Assuming a solution of the form $z = z_0 e^{i\omega t}$,

$$(-\omega^2 + i\omega\gamma + \omega_0^2)z_0 = \frac{eE_0}{m}$$

or

$$z_0 = \frac{eE_0/m}{(\omega_0^2 - \omega^2) + i\omega\gamma} = \frac{eE_0/m}{[(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2]^{\frac{1}{2}}} e^{-i\phi}$$

where $\phi = \tan^{-1} [\gamma\omega/(\omega_0^2 - \omega^2)]$. Thus the dipole moment is

$$F = P_0 \cos (\omega t - \phi)$$

where

$$P_0 = \frac{(e^2 E_0 / m)}{[(\omega_0^2 - \omega^2)^2 + \omega^2 \gamma^2]^{\frac{1}{2}}}$$

The insertion of this value of the dipole moment in the expression for Poynting's vector yields the scattered energy. This can be written conveniently in terms of the average vector for the oncoming wave, which in terms of E_0 is

$$\bar{N}_0 = \epsilon_0 c \bar{E}^2 = \frac{1}{2} \epsilon_0 c E_0^2$$

Thus \bar{N}_s , the average energy scattered per unit time per unit area in the direction θ , is, as illustrated in Fig.

1.15,

$$\begin{aligned}\bar{N}_s &= \frac{\omega^4}{32\pi^2\epsilon_0 c^3} \frac{P_0^2 \sin^2 \theta}{r^2} \\ &= \frac{r_0^2 \omega^4 \sin^2 \theta}{r^2 [(\omega_0^2 - \omega^2)^2 + \omega^2 \gamma^2]} \bar{N}_0\end{aligned}$$

where r_0 is given by Eq. (1.26).

The coefficient of \bar{N}_0/r^2 is known as the *scattering cross section*. This is a reasonable term as the coefficient represents the fraction of the energy of the primary beam scattered by the oscillator per unit solid angle in the direction θ . Alternatively, the scattering cross section, commonly written $\sigma(\theta)$, may be thought of as the area normal to the primary beam or the "window" from which the scattering system intercepts energy and redirects it as scattered radiation per unit solid angle. Thus

$$\sigma(\theta) = \frac{r_0^2 \omega^4 \sin^2 \theta}{[(\omega_0^2 - \omega^2)^2 + \omega^2 \gamma^2]} \quad (1.28)$$

The total scattering cross section comes from integrating this quantity over the unit sphere.

$$\begin{aligned}\sigma &= \int_0^\pi 2\pi \sin \theta \sigma(\theta) d\theta \\ &= \frac{8\pi r_0^2 \omega^4}{3[(\omega_0^2 - \omega^2)^2 + \omega^2 \gamma^2]}\end{aligned}$$

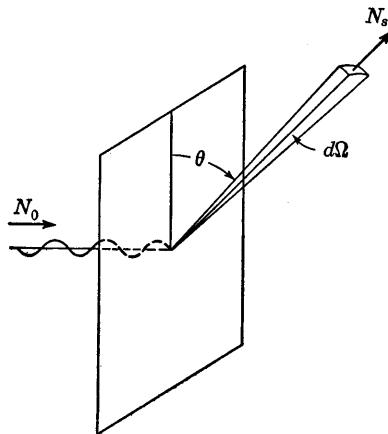


FIG. 1.15. Scattering-cross-section concepts.

This is the effective area normal to the beam from which the oscillator extracts energy from the incident radiation through a unit area and

redirects or scatters it. If there is no restoring force, as for instance for a free electron, ω_0 vanishes. Thus for a free electron

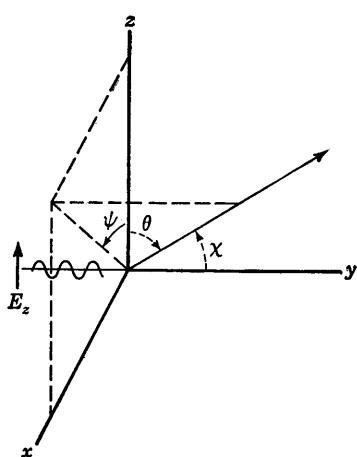


FIG. 1.16. Thomson scattering.

around the direction of propagation. From Fig. 1.16, $\cos \theta = \sin \chi \cos \psi$ and

$$\begin{aligned}\overline{\sin^2 \theta} &= \frac{\int_0^{2\pi} \sin^2 \theta d\psi}{\int_0^{2\pi} d\psi} = \frac{\int_0^{2\pi} (1 - \cos^2 \theta) d\psi}{\int_0^{2\pi} d\psi} \\ &= \frac{\int_0^{2\pi} (1 - \sin^2 \chi \cos^2 \psi) d\psi}{\int_0^{2\pi} d\psi} = 1 - \frac{1}{2} \sin^2 \chi = \frac{1}{2} (1 + \cos^2 \chi)\end{aligned}$$

Therefore, from (1.28),

$$\overline{\sigma(x)} = \frac{r_0^2 \omega^4}{2[(\omega_0^2 - \omega^2)^2 + \omega^2 \gamma^2]} (1 + \cos^2 \chi)$$

The quantity in the right-hand parentheses, which gives the angular dependence, has maxima at $\chi = 0$ and $\chi = \pi$ and a minimum at $\chi = \pi/2$. The intensity is twice as great forward and backward as in the plane normal to the radiation beam. This type of radiation scattering is known as *Thomson scattering*.

1.9. Relativistic Formulation of Electrodynamics. The equations of electricity, such as (1.17) to (1.20), may be shown to be invariant under a Lorentz transformation. Net electric charge is a constant quantity, the same when measured by any observer. Charge density transforms as does a reciprocal volume. The transformations of the field vectors and potentials can also be derived in a straightforward way. It is somewhat shorter to show, however, that all these matters are consistent

since $\gamma/\omega = 2\omega r_0/3c$ is negligible for any atomic frequency.

The angular dependence of intensity of scattered radiation is also of interest. For a polarized beam of radiation this is given by Eq. (1.28), where θ is the angle between the electric vector \mathbf{E} and the direction of observation. For an unpolarized beam and observation at an angle χ with the beam direction, the average value of $\sin^2 \theta$ must be found for χ using equal angular elements $d\psi$

with a formulation of the energy of a particle of charge e and rest mass m_0 in a region where the scalar and vector potentials are φ and \mathbf{A} , respectively. Writing H for this energy, the relativistic formulation for the conservation of energy is

$$\frac{1}{c^2} (H - e\varphi)^2 - (\mathbf{p} - e\mathbf{A})^2 = (m_0 c)^2 \quad (1.29)$$

From this it is seen that \mathbf{A} and \mathbf{p} must transform as momenta and φ as the energy E . Thus, by analogy with Eq. (1.16),

$$\begin{aligned} A_x &= A'_x, & A_y &= A'_y \\ A_z &= \frac{A'_z + \varphi' v/c^2}{\sqrt{1 - v^2/c^2}}, & \varphi &= \frac{\varphi' + v A'_z}{\sqrt{1 - v^2/c^2}} \end{aligned}$$

The electric field \mathbf{E} and magnetic induction \mathbf{B} are found in terms of \mathbf{E}' and \mathbf{B}' by taking $-(\nabla\varphi + \partial\mathbf{A}/\partial t)$ and $\nabla \times \mathbf{A}$. On forming these quantities it is seen that the presence of an electric field in one reference system contributes to a magnetic induction for an observer in motion relative to it, and similarly a magnetic induction appears to contribute to an electric field for a moving observer. The equations are found to be

$$\begin{aligned} \mathbf{B}_{\parallel} &= \mathbf{B}'_{\parallel} \\ \mathbf{B}_{\perp} &= \frac{1}{\sqrt{1 - v^2/c^2}} \left[\mathbf{B}'_{\perp} + \frac{1}{c^2} (\mathbf{v} \times \mathbf{E}')_{\perp} \right] \\ \mathbf{E}_{\parallel} &= \mathbf{E}'_{\parallel} \\ \mathbf{E}_{\perp} &= \frac{1}{\sqrt{1 - v^2/c^2}} [\mathbf{E}'_{\perp} - (\mathbf{v} \times \mathbf{B}')_{\perp}] \end{aligned}$$

where \parallel and \perp indicate parallel to and perpendicular to \mathbf{v} .

Likewise, the force as given by Eq. (1.17) can be obtained from Eq. (1.29). Since $\partial H/\partial q_x = -\dot{p}_x$, $\partial H/\partial p_x = \dot{q}_x$, and

$$H = e\varphi + c[(m_0 c)^2 + (\mathbf{p} - e\mathbf{A})^2]^{\frac{1}{2}}$$

where φ and \mathbf{A} are functions of x , y , and z ,

$$\begin{aligned} \dot{q}_x &= \dot{x} = c[(m_0 c)^2 + (\mathbf{p} - e\mathbf{A})^2]^{-\frac{1}{2}} (p_x - eA_x) \\ -\dot{p}_x &= e \frac{\partial \varphi}{\partial x} - ec[(m_0 c)^2 + (\mathbf{p} - e\mathbf{A})^2]^{-\frac{1}{2}} \left[(p_x - eA_x) \frac{\partial A_x}{\partial x} \right. \\ &\quad \left. + (p_y - eA_y) \frac{\partial A_y}{\partial x} + (p_z - eA_z) \frac{\partial A_z}{\partial x} \right] \\ &= e \frac{\partial \varphi}{\partial x} - e \left(\frac{\partial A_x}{\partial x} \dot{x} + \frac{\partial A_y}{\partial x} \dot{y} + \frac{\partial A_z}{\partial x} \dot{z} \right) \end{aligned}$$

Also the total derivative of A_x with respect to t is

$$\frac{dA_x}{dt} = \frac{\partial A_x}{\partial t} + \frac{\partial A_x}{\partial x} \dot{x} + \frac{\partial A_x}{\partial y} \dot{y} + \frac{\partial A_x}{\partial z} \dot{z}$$

Therefore, adding $-\dot{p}_x$ and $e\dot{A}_x$ and using the definitions in Sec. 1.7,

$$\begin{aligned} -\dot{p}_x + e\dot{A}_x &= e \frac{\partial \varphi}{\partial x} + e \frac{\partial A_x}{\partial t} - ej \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) + ez \left(\frac{\partial A_z}{\partial z} - \frac{\partial A_x}{\partial x} \right) \\ \frac{d}{dt} (p_x - eA_x) &= e [E_x + (\mathbf{v} \times \mathbf{B})_x] \end{aligned} \quad (1.30)$$

Thus the equation of motion is of the form of Eq. (1.17) if the force is equal to the rate of change of mechanical momentum and $\mathbf{p} - e\mathbf{A}$ is the mechanical momentum in the presence of a vector potential. Thus $\mathbf{p} = m\mathbf{v} + e\mathbf{A}$, or the generalized momentum \mathbf{p} is the sum of the mechanical momentum $m\mathbf{v}$ and what may be thought of as the electromagnetic momentum $e\mathbf{A}$. The total energy H is the mechanical energy plus the electric energy $e\varphi$. These interpretations are of course consistent with the way in which Eq. (1.29) is written, and the prediction of Eq. (1.30) justifies Eq. (1.29) as the energy equation. With the aid of the above equation all the problems of classical relativistic electrodynamics can be attacked.

PROBLEMS

- 1.1.** Show that the gravitational energy represented by the configuration of two electrons a distance apart equal to their "classical radius" ($e^2/4\pi\epsilon_0 mc^2$) is 1.962×10^{-56} joule or 1.225×10^{-37} ev. Show that for a charge-to-mass ratio of 8.607×10^{-11} coulomb kg⁻¹ for each of two identical charged mass points, the gravitational force of attraction equals the electrostatic force of repulsion.

- 1.2.** The angular momentum of a series of mass points m_i distant \mathbf{r}_i from an arbitrary origin is defined as $\mathbf{L} = \sum_i m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i$. Show that if the center of mass is at rest

the angular momentum is independent of the point of origin.

- 1.3.** Show by means of Fourier-transform pairs that the frequency spectra of the following time-pulse shapes are

$$\begin{aligned} f(t) &= e^{-t^2/2} & g(\omega) &= e^{-\omega^2/2} \\ f(t) &= 1 (-T < t < T) \\ &= 0 \text{ (otherwise)} & g(\omega) &= \sqrt{\frac{2}{\pi}} T \frac{\sin \omega T}{\omega T} \\ f(t) &= e^{-\alpha t} \sin \omega' t \ (t > 0) \\ &= 0 \ (t < 0) & g(\omega) &= \frac{1}{\sqrt{2\pi}} \frac{\omega'}{\omega'^2 - \omega^2 + \alpha^2 + 2i\omega\alpha} \end{aligned}$$

- 1.4.** From the definitions of the thermodynamic potentials prove the following equations, which are known as Maxwell's equations of thermodynamics.

$$\begin{aligned} \left(\frac{\partial T}{\partial V} \right)_S &= - \left(\frac{\partial p}{\partial S} \right)_V, & \left(\frac{\partial T}{\partial P} \right)_S &= \left(\frac{\partial V}{\partial S} \right)_P \\ \left(\frac{\partial p}{\partial T} \right)_V &= \left(\frac{\partial S}{\partial V} \right)_T, & \left(\frac{\partial V}{\partial T} \right)_P &= - \left(\frac{\partial S}{\partial p} \right)_T \end{aligned}$$

1.5. Show that the Gibbs and Helmholtz free energies can be written in terms of the enthalphy and internal energy by the following equations, known as the *Gibbs-Helmholtz* equations.

$$\Phi = x + T \left(\frac{\partial \Phi}{\partial T} \right)_v \quad \text{and} \quad \Psi = U + T \left(\frac{\partial \Psi}{\partial T} \right)_v$$

1.6. An electromagnetic oscillator is emitting a frequency ν_0 . An observer is moving with a velocity v in a direction making an angle α with the line from the oscillator to himself. Show that the frequency ν he observes is given by

$$\nu = \nu_0 \frac{1 - (v/c) \cos \alpha}{\sqrt{1 - v^2/c^2}}$$

1.7. Using the equations for the Lorentz transformation of the electric and magnetic potentials, derive the equations for the transformation of the electric field and magnetic induction.

1.8. Following the development in Sec. 1.6, assume that there is a third coordinate system designated by double primes that is moving in the x'_s direction with a velocity u as seen by an observer in the primed system of coordinates. Show that the equations of transformation from the unprimed to the doubly primed system are of the same form as Eqs. (1.13) but that the velocity v in those equations is replaced by $(u + v)(1 + uv/c^2)^{-1}$.

1.9. An observer who is at rest with respect to two equal masses compresses a weightless spring between them. He then releases the spring and observes the masses to fly apart in opposite directions. A second observer moving with respect to the first, in a direction parallel to that of the masses, observes the phenomena also. Using the relativistic definitions of mass, momentum, and energy and the equation for the addition of velocities from Prob. 1.8, show that both observers find the phenomena consistent with the conservation and transformation laws for energy and momentum.

1.10. Using the definitions for \mathbf{u} and \mathbf{A} given in the text, show that the magnetic induction due to a magnetic moment \mathbf{u} at a distant point is given by $(-1/4\pi\epsilon_0 c^2)\nabla \times [\mathbf{u} \times \nabla(1/r)]$ and the force and torque on \mathbf{u} in a magnetic induction \mathbf{B} by $\nabla(\mathbf{u} \cdot \mathbf{B})$ and $\mathbf{u} \times \mathbf{B}$, respectively.

1.11. Show that the magnetic energy represented by two dipole moments \mathbf{u}_1 and \mathbf{u}_2 separated by the vector distance \mathbf{r} is

$$U_m = \frac{1}{4\pi\epsilon_0 c^2} \left(\frac{\mathbf{u}_1 \cdot \mathbf{u}_2}{r^3} - 3 \frac{(\mathbf{u}_1 \cdot \mathbf{r})(\mathbf{u}_2 \cdot \mathbf{r})}{r^5} \right)$$

1.12. From the preceding problem find the maximum and minimum energies for a given r and variable angles of orientation of \mathbf{u}_1 . Show that the average energy for all orientations of \mathbf{u}_1 is zero. If \mathbf{u}_1 and \mathbf{u}_2 are fixed in direction, show that the average energy over all orientations of \mathbf{r} vanishes.

1.13. A particle with a charge of -1.6×10^{-19} coulomb and a mass of 9.1×10^{-31} kg rotates in a circular orbit of radius 0.53×10^{-10} m about a particle of equal and opposite charge and very great mass. Assuming the rate of rotation is such that the orbit is an equilibrium one, show that (a) the kinetic energy is equal to 2.18×10^{-18} joule or 13.6 ev, (b) this is $-\frac{1}{2}$ the potential energy of the charges, (c) the angular momentum is 1.05×10^{-34} kg-m² sec⁻¹, (d) the angular velocity is 0.413×10^{17} rad sec⁻¹.

1.14. Show that if the electron in the preceding problem radiates, it loses a fraction equal to 1.62×10^{-6} of its energy per revolution, which corresponds to radiation at

the rate of 2.33×10^{-8} watts. What is the width between half-maximum points for the curve of radiated energy as a function of frequency?

1.15. If the radiation from a charge e circulating in an orbit of radius r with an angular velocity ω is observed at a great distance \mathbf{R} in a direction making an angle ϕ with the normal to the orbit, show that the electric vector in the wave can be represented by

$$\mathbf{E} = \frac{er\omega^2}{4\pi\epsilon_0 c^2 R} (\alpha \cos \phi \cos \omega t + \beta \sin \omega t)$$

where α is a unit vector normal to \mathbf{R} lying in the plane containing \mathbf{R} and the orbit normal, and β is a unit vector perpendicular to \mathbf{R} and α . Describe the nature of the polarization of the wave as a function of ϕ . Using the values of e , r , and ω from Prob. 1.13, show that the average power per unit area radiated in the direction of the normal to the orbit at a distance of 1 m is 5.57×10^{-9} watt m^{-2} .

1.16. An electromagnetic wave is incident upon an atomic oscillator (a bound electron) having a natural angular velocity of $\omega_0 = 0.41 \times 10^{17}$ rad sec $^{-1}$. Assuming the only damping is that due to radiation, show that the total cross section for scattering is $(3/2\pi)\lambda_0^2$, where λ_0 is the wavelength of a wave of angular frequency ω_0 , when the incident radiation frequency is equal to the oscillator natural frequency. Show that this is 1.5×10^{13} times the cross section for scattering by a free electron. How does this scattering cross section compare with the area of the orbit in Prob. 1.13? What values does the cross section approach for very high and very low frequencies of the incident wave?

1.17. Show that in the nonrelativistic approximation Eq. (1.29) becomes

$$H \cong m_0 c^2 + \frac{(p - eA)^2}{2m_0} + e\phi \cong m_0 c^2 + \frac{1}{2}m_0 v^2 + e\phi$$

1.18. Consider nonrelativistically the radiation from a charge e rotating with an angular velocity ω in a circular orbit of radius l about the origin in the xy plane. From the equation for the vector potential, $\mathbf{A} = \frac{1}{4\pi\epsilon_0 c^2} \frac{e\mathbf{v}}{r}$ where \mathbf{v} and r are evaluated at a time $t = r/c$, neglecting $1/r$, and considering $l\omega/c$ to be small, show that

$$\mathbf{A} = \frac{e\omega l}{4\pi\epsilon_0 c^2} \frac{1}{R} \left[i \left(-\sin \omega T - \frac{\omega \delta R}{c} \cos \omega T \right) + j \left(\cos \omega T - \frac{\omega \delta R}{c} \sin \omega T \right) \right]$$

where R is the radius vector from the origin,

$$\delta R = l[\cos \alpha_1 \cos \omega(T - R/c) + \cos \alpha_2 \sin \omega(T - R/c)],$$

$T = t - R/c$, α_1 and α_2 are the angles between R and the x and y axes, respectively, and i and j are unit vectors along these axes. Using this value of \mathbf{A} , calculate \mathbf{B} in the radiation field where $R \gg l$ for the two extensions of this configuration for which (a) a charge $+e$ and $-e$ at opposite ends of a diameter rotate in the circle of radius l , and (b) two charges each of $+e$ at opposite ends of a diameter rotate in the circle of radius l . The first is an electric dipole and the second is called a magnetic dipole. Show that the average value of the ratio of the power radiated by the magnetic dipole to that radiated by the electric dipole is $[(2\pi l/\lambda) \sin \alpha_3]^2$, where λ is the wavelength and α_3 is the angle between R and the z axis.

1.19. The close juxtaposition of n dipoles leads to a configuration known as a 2^n -tuple pole. An example of a quadrupole, for instance, would be two dipoles a distance apart comparable to the charge separation of each dipole. Consider an oscillating quadrupole formed by two identical, oppositely phased collinear oscillating

dipoles with their centers separated by a small distance δl . Using Eq. (1.21) show that the magnitude of the electric or magnetic vector in the radiation field from such a combination is $2\pi \sin \theta \delta l/\lambda$ times the field of a dipole, where θ is the angle with the dipole axis and $\nu = c/\lambda$ is the frequency of oscillation. Hence, neglecting angular factors, show that the rate of radiation of energy from a 2^n -tuple pole is of the order of $(\delta l/\lambda)^{2(n-1)}$ times the rate of radiation of energy from a dipole.

GENERAL REFERENCES

- Bergmann, P.: "Introduction to the Theory of Relativity," Prentice-Hall, Inc., New York, 1946.
- Bridgman, P. W.: "The Logic of Modern Physics," The Macmillan Company, New York, 1927.
- Bridgman, P. W.: "The Nature of Physical Theory," Princeton University Press, Princeton, N.J., 1936.
- Einstein, A., and L. Infeld: "The Evolution of Physics," Simon and Schuster, Inc., New York, 1938.
- Epstein, P. S.: "Textbook of Thermodynamics," chaps. 2, 4, 6, 17-20, John Wiley & Sons, Inc., New York, 1937.
- Goldstein, H.: "Classical Mechanics," chaps. 2, 6, 7, Addison-Wesley Publishing Company, Cambridge, Mass., 1951.
- Harnwell, G. P.: "Principles of Electricity and Electromagnetism," 2d ed., chap. 16, McGraw-Hill Book Company, Inc., New York, 1949.
- Houston, W. V.: "Principles of Mathematical Physics," 2d ed., chaps. 6, 10, 15, 16, McGraw-Hill Book Company, Inc., New York, 1948.
- Joos, G.: "Theoretical Physics," chaps. 2, 6, 10, 17, 26, 29, G. E. Stechert & Company, New York, 1934.
- Mees, C. E. K.: "The Path of Science," John Wiley & Sons, Inc., New York, 1946.
- Page, L.: "Introduction to Theoretical Physics," 2d ed., chaps. 4, 7, 12, D. Van Nostrand Company, Inc., New York, 1935.
- Page, L., and N. I. Adams: "Principles of Electricity," 2d ed., chap. 16, D. Van Nostrand Company, Inc., New York, 1949.
- Robertson, H. P.: *Rev. Mod. Phys.*, **21**, 378 (1949) (relativity).
- Slater, J. C., and N. H. Frank, "Mechanics," chap. 4, McGraw-Hill Book Company, Inc., New York, 1947.
- Weisskopf, V. F.: *Rev. Mod. Phys.*, **21**, 305 (1949) (theory of the electron).
- Whitehead, A. N.: "The Principles of Natural Knowledge," Cambridge University Press, London, 1925.
- Zemansky, M. W.: "Heat and Thermodynamics," 3d ed., chaps. 4, 7, 10, 15, McGraw-Hill Book Company, Inc., New York, 1951.

CHAPTER 2

ATOMIC NATURE OF MATTER AND RADIATION

2.1. Early Evidence for Atomicity. Many of our current physical concepts appear to have been evolved in the dawn of Greek intellectual activity and bear the marks of the inspired speculation of that era. Greek philosophers propounded the concept of atomicity together with what was thought of as its antithesis, continuity. Continuity may be roughly characterized as representing indefinite divisibility, and atomicity by its opposite. Atomicity was possibly first proposed by Leucippus as a means out of Zeno's paradox of the hare and the tortoise. Though an interesting idea in itself, it was actually irrelevant to that paradox since the basic difficulty of Zeno was an uncritical acceptance of his method of approaching a limit. The atomicity of matter and space was advocated by Democritus of Abdera about 400 B.C., and the atomicity of all nature was taught by Epicurus about 100 years later. The atomic concept was due largely to these early philosophers, but otherwise their scientific contribution to the subject of physics was inconsiderable since they had little appreciation of the role of observation and experiment in the discipline of theory.

Cogent arguments for the atomicity of matter were not actually educed till about 200 years ago, when natural philosophers noted the fact that the existence of recognizable substances recoverable from homogeneous mixtures was difficult to reconcile with any theory of the continuity of matter. Daniel Bernoulli, who was a scientist of great insight, described in 1738 how the behavior of a gas—as the pressure and temperature were changed—could be explained on the hypothesis that the gas consisted of a large number of infinitesimal particles in rapid random motion. This is the point of view of the present kinetic theory of gases. The most influential evidence, however, for the atomic hypothesis probably came from the two early chemical laws of definite and multiple proportions which are generally associated with the names of J. L. Prout and J. B. Richter in the latter part of the eighteenth century. These state, respectively: (1) The proportions in which two elements combine cannot vary continuously. (2) If two elements combine together in more than one way, the masses of one that combine with a given mass of the other

are in simple ratios to one another. These laws are summarized succinctly in the present common notation for chemical formulas. But it was not until 1808 that John Dalton marshaled and presented the evidence sufficiently convincingly to sway scientific opinion to the acceptance of the atomic theory.

The early observations on the properties of gases were particularly influential in the development of atomic theory. Robert Boyle in 1661 showed that at a constant temperature the density of a gas is proportional to the pressure. It was not until 1802, when more had been learned about heat and certain imperfect temperature scales had been developed, that Gay-Lussac showed that at a constant pressure the density of a gas varies inversely as the temperature. Gay-Lussac also showed that the volumes of gas that appear or disappear in a reaction are in simple ratios to each other. The gas laws may be stated by saying that the product of the temperature, on a properly chosen scale, and the density is proportional to the pressure. The synthesizing step that was taken by Avogadro in 1811 was essentially the observation that if the masses of gas samples are properly chosen the constant that occurs in the gas law, when written as an equation, is a universal constant applying to all gases. The mass of gas to be chosen is proportional to what is now known as the molecular weight, and Avogadro recognized this mass as being proportional to the weight of a molecule of the gas. From the laws of combining proportions and the universal gas law Avogadro drew the logical conclusion that equal volumes of gases under the same conditions of temperature and pressure contain equal numbers of molecules. This is of course in accord also with Gay-Lussac's observations on gas reactions.

On this conceptual foundation the chemists of the following generation proceeded rapidly to determine the atomic weights of the known elements by comparing combining weights. Since the number of atoms of an element in a given mass was not known to Avogadro or the later chemists, there was no uniquely indicated choice for the atomic weight and all that was significant was the ratio of combining weights. Hence an arbitrary choice for a fiducial element mass had to be made, and, because of the many combinations into which it entered, oxygen was chosen and assigned the arbitrary atomic weight of 16. On this basis the smallest atomic weight, that of hydrogen, became about 1 and the other known elements could be arranged serially in terms of atomic weight up through uranium with an atomic weight of about 238.

In addition to this serial classification, similarities in chemical properties such as those between fluorine, chlorine, and bromine and between lithium, sodium, and potassium suggested a periodic recurrence of related attributes. On this evidence D. I. Mendeleev proposed in 1869 the

scheme of ordering which we know as the periodic table. This may be described by saying that the properties of elements are periodic functions of their atomic weights. Much still remained to be done before the periodic table assumed its present form. However, this step was very stimulating in suggesting hitherto unknown elements indicated by gaps in the table. After the existence of isotopes was discovered and the more fundamental role of atomic number appreciated, the table was of further value in suggesting the nature of electronic structures characteristic of the successive atoms in the sequence of the elements.

2.2. Avogadro's Number. Two of the first items of interest in regard to atoms are their sizes and masses. The size or extent of an atom is not a well-defined concept, but there was much early interest in it, and an upper limit is clearly given by the lower limits on particle sizes and film thicknesses. These are, however, relatively crude approximations since such measurements are limited by the wavelength of visible light to linear dimensions of the order of 10^{-6} to 10^{-7} m. These lengths are about 1000 times as great as the order of magnitude of atomic dimensions defined by some significant structural feature. A more interesting and significant problem is that of determining atomic masses. This problem can be best approached through a determination of the number of atoms which constitute some mass of easily measurable magnitude such as a *mole*, which is defined as the mass in grams numerically equivalent to an atomic or molecular weight. A *kilomole* is a mass in kilograms numerically equal to the molecular weight, and it contains a thousand times as many molecules as a mole. The number of molecules in a mole is known as *Avogadro's number*.

Certain early experiments for the measurement of Avogadro's number N were of a type which verified the general correctness of kinetic-theory concepts and yielded approximate values of N , reliable to within a few per cent. One of these was the measurement by Perrin¹ of the number of colloidal mastic particles in a liquid suspension as a function of height. On the assumption of the applicability of statistical mechanics to the distribution of such suspended particles in the earth's gravitational field (Sec. 6.3), the ratio of the numbers per unit volume at two heights, say h_1 and h_2 ($h_2 > h_1$), is given by

$$\frac{n_2}{n_1} = e^{-(m'g/kT)(h_2-h_1)} \quad (2.1)$$

T is the absolute temperature of the fluid, g is the gravitational constant, and m' is the effective mass. The effective mass is $v(\rho - \rho')$, where v is

¹ J. Perrin, "Atoms," Constable & Co., Ltd., London, 1923.

the particle volume and ρ and ρ' are the densities of the particle material and fluid, respectively. k is Boltzmann's constant, which is the gas constant per mole divided by Avogadro's number. The gas constant $R = 8.316 \text{ joules deg}^{-1} \text{ mole}^{-1}$ is known, so $N = R/k$ can be found by measuring all the other quantities in Eq. (2.1). The precision of this method is not great.

A second approach to the measurement of an atomic mass or Avogadro's number is through the quantitative analysis of Brownian motion. This is the residual thermal motion exhibited by any system in thermal equilibrium with its surroundings. It was first observed by the botanist Robert Brown over a century ago when he remarked the rapid random motion of plant spores suspended in a liquid under his microscope. More carefully controlled and systematic observations by Perrin on the motion of gamboge particles in a water suspension provided the data to which an analysis by Einstein could be applied. The value of the Boltzmann constant is determined from a knowledge of the temperature and the mean-square displacement experienced by a particle during a sequence of equal time intervals. If x is the observed displacement along some selected coordinate observed during an interval τ , it may be shown by an analysis similar to that leading to Eq. (6.15) that the mean-square value of x (see Fig. 2.1) is given by

$$\overline{x^2} = \frac{2kT}{\alpha} \tau$$

Here α is the viscous retarding force per unit velocity that slows down the particle when moving through the suspending fluid. If the coefficient of viscosity is η and the particle is spherical of radius a , α is given approximately by $6\pi\eta a$. Thus from a series of observations of τ and x a fairly reliable value of $2kT/\alpha$ can be obtained. Assuming $\alpha = 6\pi\eta a$, that η and a are known, and that the temperature is measured, a value for k is found. The constant k being known, N , which is equal to R/k , is found from the gas constant R . The precision of this type of measurement is not great; it is comparable with the immediately preceding one dealing with the distribution of particle density with height in a suspension. The measurements yield values of the order of $6.1 \times 10^{23} \text{ mole}^{-1}$ for N .

Avogadro's number can be determined most accurately by utilizing the regularity of structure characteristic of a crystal and making a care-

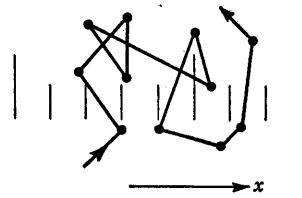


FIG. 2.1. Brownian displacements.

ful absolute measurement of the lattice spacing (Sec. 7.6).¹ From a measurement of this spacing the volume occupied by a molecule in a perfect crystal can be obtained. Multiplication of this volume by the density of the crystal gives the mass of a molecule. This last is the quantity of primary interest. From it can be obtained the masses of the constituent atoms since their relative weights are well known from chemical and other methods. The quotient of the molecular weight and the mass of the molecule yields Avogadro's number.

The experimental steps in the precise determination of the mass of an atom of unit atomic weight or Avogadro's number furnish an interesting illustration of an elementary atomic measurement. First the wavelength of a suitable spectrum line free of complex structure is measured

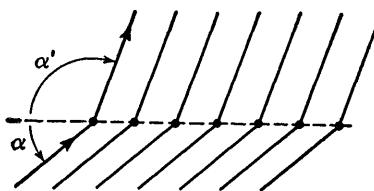


FIG. 2.2. Plane grating diffraction.

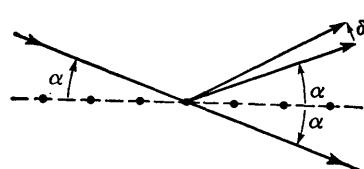


FIG. 2.3. Plane grating diffraction at grazing incidence.

by means of an interferometer in terms of the standard meter. Let the numerical value so determined be indicated by λ_l . Then the angle of occurrence of, say, the p th-order diffraction maximum for this light when incident normally on a diffraction grating is measured. Using the geometry of Fig. 2.2, the general diffraction-maximum condition for a grating is

$$p\lambda_l = d_g(\cos \alpha + \cos \alpha') \quad (2.2)$$

Assuming normal incidence, $\cos \alpha$ equals zero and hence, since p and λ_l are known, d_g is given by $d_g = p\lambda_l/\cos \alpha'$. The type of grating used is such that the spacing d_g between the rulings is not very much larger than λ_l , so that the condition may be fulfilled for small values of p and values of $\cos \alpha'$ of the order of unity. In the third step the same grating is used at grazing incidence, as in Fig. 2.3, to determine with approximately equal accuracy the wavelength λ_x of a suitably well-defined X-ray line which is of the order of 10^{-4} as great as λ_l . By using a sufficiently great path length the angles between the transmitted and reflected rays, 2α , and between the reflected and first-order diffracted rays, $\alpha' - \alpha = \delta$, can be measured with high precision. Using the diffraction condition and the approximations for small angles,

¹ See R. T. Birge, *Am. J. Phys.*, **13**, 63 (1945) for detailed references.

$$\begin{aligned}\lambda_x &= d_g[\cos \alpha - \cos (\alpha + \delta)] \\ &= d_g(\cos \alpha - \cos \alpha \cos \delta + \sin \alpha \sin \delta) \\ &= d_g\left(\frac{\delta^2}{2} \cos \alpha + \delta \sin \alpha\right) \\ &= d_g\left(\frac{\delta^2}{2} + \alpha \delta\right)\end{aligned}$$

to the second order of small quantities. Thus, using $d_g = p\lambda_l/\cos \alpha'$,

$$\lambda_x = \lambda_l \frac{p}{\cos \alpha'} \left(\frac{\delta^2}{2} + \delta \alpha \right) \quad (2.3)$$

By the above procedures the wavelength of the X radiation, which is very much smaller than the grating space, can be measured in terms of λ_l with high precision.

The final step is then to measure the grating space of a suitable crystal by means of λ_x and hence in terms of λ_l . The crystal must be a pure, permanent, rigid structure as free as possible of fissures or other defects that would introduce errors. One of the most suitable crystals for the purpose has been found to be calcite (CaCO_3). X-ray diffraction provides the most important technique for determining not only the crystal class and the dimensions of the unit lattice cell, but also the electron distribution within the crystal and the size and nature of any imperfections in the crystal specimen. The perfect crystal, which for the purpose of this discussion will be assumed to be used, is characterized primarily by a regular space lattice of scattering centers. These scattering centers are the groups of electrons closely associated with the atomic nuclei which are arranged in a highly regular repetitive pattern. The pattern is characterized by the fact that there are three vectors, which do not lie in a plane, along any one of which the crystal may be assumed to be displaced by a certain characteristic unit length with the result that the identical pattern of scattering centers recurs. Let the three vectors describing the crystal be of length l_1 , l_2 , and l_3 with direction cosines (a_1, b_1, c_1) , (a_2, b_2, c_2) , and (a_3, b_3, c_3) , respectively, as shown in Fig. 2.4. Then if the direction of the incident radiation is given by the direction cosines $(\alpha_1, \alpha_2, \alpha_3)$ and the direction of the diffracted radiation is given by

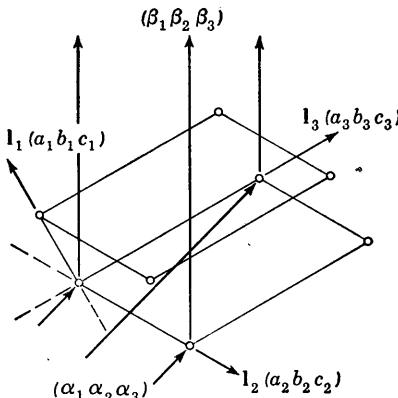


FIG. 2.4. Three-dimensional diffraction.

the direction cosines ($\beta_1, \beta_2, \beta_3$), the condition for constructive interference between a wave scattered by one atom and that scattered by a second atom at the vector separation l_1 from the first is that the difference between the projection of l_1 along the incident direction and along the diffracted direction shall be an integral number of wavelengths:

$$n_1\lambda = l_1[(a_1\alpha_1 + b_1\alpha_2 + c_1\alpha_3) - (a_1\beta_1 + b_1\beta_2 + c_1\beta_3)]$$

Constructive interference for scattering by atoms displaced l_2 and l_3 from the first leads to analogous expressions, and if all three of these conditions are satisfied it is evident that the waves scattered by all the lattice points will interfere constructively. The intensity of the scattered or diffracted wave in the direction ($\beta_1, \beta_2, \beta_3$) will then be a maximum. It is the necessity for satisfying three diffraction conditions to obtain a maximum of intensity instead of the single condition of Eq. (2.2) for an optical grating that differentiates crystal diffraction from ordinary grating diffraction. This will be seen to lead to a condition between the wavelength, crystal parameters, and angle between incident and diffracted rays for constructive interference to take place at all.

Writing the three equations analogous to the one above and factoring out the length and direction cosines of the lattice vectors,

$$\begin{aligned}\frac{n_1\lambda}{l_1} &= a_1(\alpha_1 - \beta_1) + b_1(\alpha_2 - \beta_2) + c_1(\alpha_3 - \beta_3) \\ \frac{n_2\lambda}{l_2} &= a_2(\alpha_1 - \beta_1) + b_2(\alpha_2 - \beta_2) + c_2(\alpha_3 - \beta_3) \\ \frac{n_3\lambda}{l_3} &= a_3(\alpha_1 - \beta_1) + b_3(\alpha_2 - \beta_2) + c_3(\alpha_3 - \beta_3)\end{aligned}$$

For the angles determined by these equations at the particular wavelength, it is as if the electromagnetic disturbance were unable to enter the crystal and hence is deflected by its presence. These equations are a little cumbersome and they may be put into a more convenient form by solving explicitly for the quantities $(\alpha_1 - \beta_1)/\lambda$, $(\alpha_2 - \beta_2)/\lambda$, and $(\alpha_3 - \beta_3)/\lambda$ in terms of n_1/l_1 , n_2/l_2 , and n_3/l_3 . Using the method of determinants, the new equations are immediately found to be

$$\begin{aligned}\frac{(\alpha_1 - \beta_1)}{\lambda} &= a'_1 \frac{n_1}{l_1} + a'_2 \frac{n_2}{l_2} + a'_3 \frac{n_3}{l_3} \\ \frac{(\alpha_2 - \beta_2)}{\lambda} &= b'_1 \frac{n_1}{l_1} + b'_2 \frac{n_2}{l_2} + b'_3 \frac{n_3}{l_3} \\ \frac{(\alpha_3 - \beta_3)}{\lambda} &= c'_1 \frac{n_1}{l_1} + c'_2 \frac{n_2}{l_2} + c'_3 \frac{n_3}{l_3}\end{aligned}\quad (2.4)$$

where the primes denote the cofactors of the letters in the determinant of the coefficients, that is, $a'_1 = (b_2c_3 - b_3c_2)/(\text{determinant of coefficients})$, etc. If the equations are then squared and added together, the left-hand side becomes

$$\begin{aligned} \frac{1}{\lambda^2} [(\alpha_1^2 + \alpha_2^2 + \alpha_3^2) - 2(\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3) + (\beta_1^2 + \beta_2^2 + \beta_3^2)] \\ = \frac{2(1 - \cos \phi)}{\lambda^2} \\ = \frac{4 \sin^2(\phi/2)}{\lambda^2} \end{aligned}$$

where ϕ is the angle between the incident and diffracted radiation. The sum of the squares of the right-hand side is more involved, but it may be shown to be the reciprocal of the square of the spacing between the family of parallel planes which intersect the three vector axes \mathbf{l}_1 , \mathbf{l}_2 , and \mathbf{l}_3 at points given by $s\mathbf{l}_1/n_1$, $s\mathbf{l}_2/n_2$, and $s\mathbf{l}_3/n_3$, respectively, s being the serial integer designating the particular plane of the family. Such families of planes pass through all the atoms or molecules in the crystal. Calling this spacing $d(n)$, the equation resulting from squaring and adding is

$$\lambda = 2 d(n) \sin \frac{\phi}{2} \quad (2.5)$$

This is the Bragg law for diffraction by a crystal lattice, illustrated in Fig. 2.5. The planes generally used for calcite are those parallel to a cleavage face for which two of the n 's are zero; and for the first-order diffraction maximum the third n is unity. Also, for calcite the l 's are all equal, which simplifies the calculation of d_c in terms of l and the crystal angles.

There is, however, one further consideration before Eqs. (2.3) and (2.5) can be used to eliminate the X-ray wavelength and determine d_c in terms of λ_i and the various angles. The index of refraction of a crystal for X rays is not unity, and hence the wavelength in the crystal [Eq. (2.5)] is not the same as that in free space [λ_x of Eq. (2.3)]. The index is measured accurately by measuring the angle of total reflexion and is found to be slightly less than unity, i.e., total reflexion, preventing radiation from entering the crystal, takes place at a small glancing angle. If i_0 is the critical angle, the index of refraction is by definition $\sin i_0$. The wavelength outside and inside the crystal and the angles outside

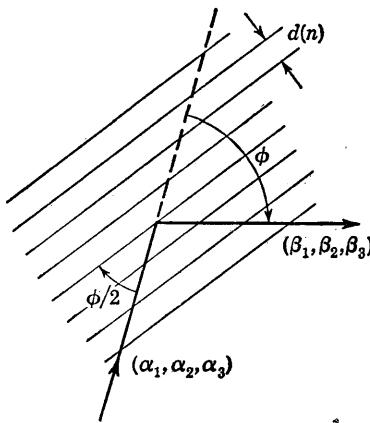


FIG. 2.5. Bragg plane concept.

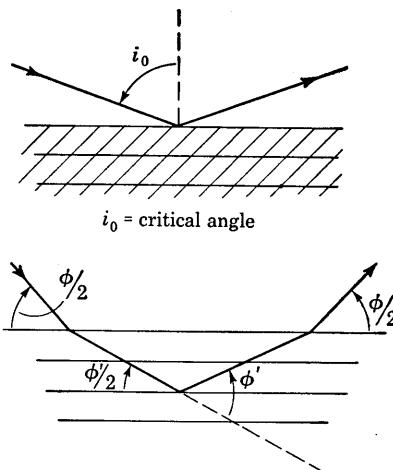


FIG. 2.6. Internal and external diffraction angles.

and inside, as shown in Fig. 2.6, are related to the index of refraction by

$$\sin i_0 = \frac{\lambda}{\lambda'} = \frac{\cos (\phi/2)}{\cos (\phi'/2)}$$

Using Eq. (2.5) in the form applicable to the interior of the crystal, $\lambda' = 2d_c \sin (\phi'/2)$, and eliminating λ' and ϕ' ,

$$\lambda = 2d_c \sin \frac{\phi}{2} \left[1 - \left(\frac{\cos i_0}{\cos (\phi/2)} \right)^2 \right]^{\frac{1}{2}}$$

where λ is now the same as λ_x of Eq. (2.3) and ϕ is the external diffraction angle. The final result of all these steps is then

$$\lambda_x = \lambda_i \frac{p}{\cos \alpha'} \left(\frac{\delta^2}{2} + \delta \alpha \right) = 2d_c \sin \frac{\phi}{2} \left\{ 1 - \left[\frac{\cos i_0}{\sin (\phi/2)} \right]^2 \right\}^{\frac{1}{2}} \quad (2.6)$$

Using this relationship, d_c is determined by an X-ray crystal spectrometer in terms of λ_i , the various angles, and the index p . The best present experimental¹ value of this constant is (at 20°C)

$$d_c = (3.03567 \pm 0.00005) \times 10^{-10} \text{ m}$$

It is then necessary to find the volume of the unit rhombohedral cell of calcite in terms of d_c and the rhombohedral angle of calcite. This can be done most easily by means of the direction cosines of the three edges. Let **OC** of Fig. 2.7 correspond with the y axis, **OA** lie in the xy plane, and r be the rhombohedral angle between **OA** and **OB**, **OB** and **OC**, and **OA** and **OC**. Then the direction cosines are

	x	y	z
OA =	$\sin r$	$\cos r$	0
OB =	p	$\cos r$	q
OC =	0	1	0

p is determined by the fact that the angle between **OA** and **OB** is r , that is, $p \sin r + \cos^2 r = \cos r$, and q from the fact that the sum of the squares of the direction cosines is unity.

$$p = \cot r (1 - \cos r)$$

$$q = [\sin^2 r - \cot^2 r (1 - \cos r)^2]^{\frac{1}{2}}$$

¹ J. A. Bearden, *J. Appl. Phys.*, **12**, 395 (1941); R. T. Birge, *Am. J. Phys.*, **13**, 63 (1945).

The volume is then $OA \times OC \cdot OB$, or l^3 times the determinant above:

$$V = l^3 q \sin r = l^3 (1 - 3 \cos^2 r + 2 \cos^3 r)^{\frac{1}{2}}$$

and, since from the figure $d_c = (l/2)q$,

$$V = 8d_c^3 \frac{\sin^3 r}{1 - 3 \cos^2 r + 2 \cos^3 r} = \frac{8d_c^3(1 + \cos r)^2}{(1 + 2 \cos r) \sin r}$$

The angle r for calcite is found to be $101^\circ 55'$. It is seen from Fig. 2.7 that the cell contains four complete CaCO_3 molecules, so the volume occupied by a molecule is $V/4$, and the mass of a calcite molecule is $\rho V/4$, where ρ is the density. Avogadro's number is then the molecular weight divided by the mass of a molecule, or

$$N = \frac{4M}{\rho V}$$

The density of CaCO_3 has been determined as $2.71030 \pm 0.00003 \text{ g cm}^{-3}$ at 20°C .

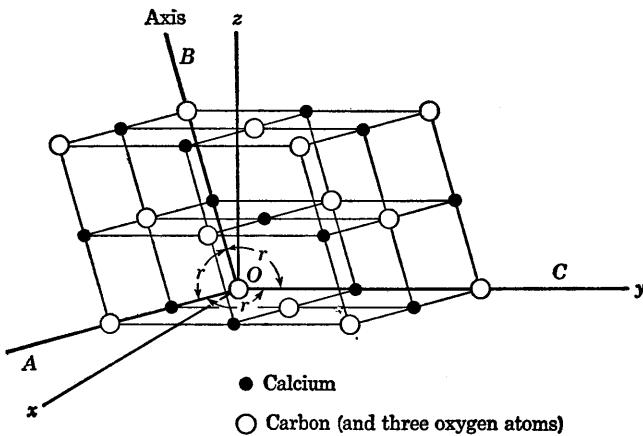


FIG. 2.7. Schematic calcite crystal.

A remark must be made on M . The chemical scale of atomic weights assumes the atomic weight of the natural mixture of oxygen isotopes in the air to be 16. Atmospheric oxygen consists approximately of 2494 parts O^{16} , 1 part O^{17} , and 5 parts O^{18} . These ratios do not remain constant owing to the differential concentration that takes place during chemical or physical processes. Therefore it is more satisfactory to define O^{16} as having the atomic weight of 16. This is known as the physical scale. Taking the isotopic weight from mass-spectrographic

data, $O^{16} = 16$, $O^{17} = 17.0045$, $O^{18} = 18.00487$; the atomic weight of atmospheric oxygen on the physical scale is 16.00433 ± 0.00001 . The ratio of the scales is 1.000272 ± 0.000005 . We shall use the physical scale. An analysis¹ of the most precise experimental determinations of N gives a value of $(6.02502 \pm 0.00020) \times 10^{23}$ molecules per mole (physical scale). A least-squares adjustment of the primary atomic constants gives a slightly smaller value $(6.02472 \pm 0.00036) \times 10^{23}$. This important basic quantity is the number of molecules in that mass of any pure substance which is equal numerically in grams to the molecular weight. The reciprocal of N is the mass of an atom of unit atomic weight M_1 . On the physical scale, this is

$$M_1 = (1.65963 \pm 0.00003) \times 10^{-27} \text{ kg} \quad (\text{physical})$$

From the above discussion it is seen that these very fundamental atomic constants are known independently of any others with an accuracy of about 1 part in 10^4 . The numerical probable errors given above are less than this because the present best values of these constants are arrived at by taking into account other related data of comparable precision as described in the references cited.

2.3. Charges and Masses of Atomic Components. Atoms are electrically neutral as is evidenced by the normally neutral condition of large pieces of matter which contain many identical atoms. However, atomic constituents are capable of participating in the process known as electric ionic conduction. In 1834 Michael Faraday published the results of his observations on the conduction of electricity through solutions. These experiments showed that the passage of equal quantities of electric charge through a solution resulted in the deposition of equal quantities of an element at an electrode and that the passage of a given charge through two solutions in series resulted in the deposition of equal numbers of gram equivalents of different elements. A gram equivalent is a mole divided by the chemical valence, or combining ratio, of an element. These experiments can be most simply interpreted if one assumes that a certain characteristic charge is associated with each atomic or molecular fragment in solution and that the valence is the number of elementary charges associated with the atomic or molecular fragment. A careful measurement of the mass of an element deposited as a result of the passage of a certain charge can be used, on the basis of Faraday's views and in conjunction with Avogadro's number, to determine the magnitude of the elementary characteristic atomic charge. Such measurements are capable of considerable precision; in fact for many years the international unit of electric charge was based upon electrolytic measurements. Now the absolute practical unit of charge is otherwise defined, and precision

¹ J. W. M. DuMond and E. R. Cohen, *Rev. Mod. Phys.*, **25**, 691 (1953).

experiments show that the passage of 1 coulomb is accompanied by the deposition of 0.0111818 ± 0.0000008 g of silver from solution. A precise determination of the atomic weight of silver on the physical scale yields for its atomic weight 107.914 ± 0.002 . The quotient of these, which is the number of coulombs that must pass through a solution to deposit a mole of an element, is known as the *faraday*. The techniques of quantitative electrolysis are difficult, and the values of this constant as derived from different electrolytic experiments have not yet attained a satisfactory consistency. DuMond and Cohen¹ quote the best present experimental values on the physical scale as

$$\begin{aligned} 1 \text{ faraday} &= 96,521.5 \pm 1.3 \text{ coulombs g-equiv}^{-1} && (\text{iodine}) \\ &= 96,512.9 \pm 1.9 \text{ coulombs g-equiv}^{-1} && (\text{silver}) \end{aligned}$$

This constant is associated with other atomic constants, and the least-squares adjusted value calculated by DuMond and Cohen is

$$1 \text{ faraday} = 96,520.1 \pm 2.5 \text{ coulombs g-equiv}^{-1}$$

If each univalent atom has one elementary atomic charge associated with it, the magnitude of this charge would be about 1.602×10^{-19} coulomb. The best present weighted mean of experimental values for this quantity is

$$e = \frac{\text{faraday}}{N} = (1.60207 \pm 0.00007) \times 10^{-19} \text{ coulomb}$$

Although the electrolytic method has until recently been the most precise one for determining e , techniques have now been developed for making precise absolute measurements of the ratio of the charge carried by an atom from which an electron has been removed to the mass of this atom. If the mass of the atom is known from a knowledge of N and the atomic weight, an accurate value for the unit atomic charge can be obtained. An experiment of this type will be described later in this section.

The nature of the atomic process whereby atoms become endowed with a net electric charge is not explained by the derivation of the value of the elementary atomic charge from Avogadro's number and the faraday. For instance, a unique elementary electric charge need not be assumed for accord with the electrolytic experiments. The value of e could be the average of atomic charges which were not all the same but simply occurred in equal proportions in any large sample of atoms of an element. However, a number of other experiments have been performed which lend general support to the concept of a universal atomic unit of electric charge. It was found by early workers that drops of liquid leaving an atomizer were, in general, electrically charged and that the charge per

¹ J. W. M. DuMond and E. R. Cohen, *Rev. Mod. Phys.*, **25**, 691 (1953).

drop was of the order of magnitude of e given above. Millikan,¹ and later Hopper and Laby,² studied the motion of a microscopic liquid droplet as it rose and fell under the influence of gravity and electric fields while free to gain or lose charge from weakly ionized air surrounding it. These experiments showed that the smallest amount by which the charge on the droplet varied was equal to e within their limits of experimental error. Thus all the charges on the drops could be represented by ne , where n was a positive or negative integer. The charge e thus assumes the position of a natural unit. No smaller subdivisions of charge than this are found in nature. Many other lines of evidence from photo-electricity, thermionics, and the internal consistency of atomic theory confirm the important fact that electric charge exists only in this universal elementary atomic unit.

Investigations of the conduction of electricity through gases first led to an understanding of the nature of the charged atomic fragments. The early experiments in electrostatics were all performed in air, and investigators during the early part of the preceding century were aware of the leakage of electric charge through air. Sir William Crookes observed certain of the phenomena associated with the passage of ions through partially evacuated tubes and, in a rare flash of insight, in 1886 proposed the existence of atoms that are chemically similar but differ in mass. Johnstone Stoney in 1891 proposed the name electron for the quantity of electricity necessary to deposit one univalent atom in electrolysis. However, the most enlightening series of experiments was that carried out by Sir J. J. Thomson³ and his pupils and colleagues at the Cavendish Laboratory, beginning about 1895, which was the same year in which X rays were discovered by Konrad Roentgen. Thomson and his associates, in a long and careful series of experiments, evolved the concepts of positive and negative ions, or electric carriers, in a gas in the conducting condition. By measurements on the motion of these ions in the presence of both electric and magnetic fields, it was shown that one type of entity, identified by its charge-to-mass ratio, appeared to be universal in that this same charge-to-mass ratio was found among electric carriers in all gases. Other types of carriers of electricity had the opposite sign of charge-to-mass ratio; these ratios were smaller in magnitude, and their values were characteristic of the particular gas through which the conduction took place. These early experiments were confirmed by many other investigators, and the particle with the universal negative charge-to-mass ratio received the name *electron*. It was observed in thermionic and photo-

¹ R. A. Millikan, *Phys. Rev.*, **29**, 560 (1909); **32**, 349 (1911); **2**, 109 (1913).

² V. P. Hopper and T. H. Laby, *Proc. Roy. Soc. (London)*, **178**, 243 (1934).

³ J. J. Thomson and G. P. Thomson, "Conduction of Electricity through Gases," Cambridge University Press, London, 1928.

electric phenomena and was soon recognized as a common constituent of all atomic systems. The particles with the smaller, positive charge-to-mass ratios were identified with the residual atomic systems from which electrons had been removed, and the fact that more than one charge-to-mass ratio was observed in a single gas confirmed Crooke's concept of *isotopes*, which are atoms of different masses but of a single chemical species.

The general equation of motion of a charge e is Eq. (1.17), and from it may be determined the trajectories of charged particles in the presence of constant or changing fields. Consider the change in momentum during a small time interval δt . Using the relation $\mathbf{E} = -\partial \mathbf{A}/\partial t - (\nabla \varphi)$ and integrating Eq. (1.17) or Eq. (1.30) over δt ,

$$\mathbf{F} \delta t = \delta \mathbf{p} = -e \frac{\partial \mathbf{A}}{\partial t} \delta t - e(\nabla \varphi) \delta t + \frac{e}{m} \mathbf{p} \times \mathbf{B} \delta t$$

The first term on the right is the impulse due to the time rate of change of \mathbf{A} , the second term is the impulse from the electric field due to the electric potential, and the third term is the impulse due to motion through the constant magnetic induction \mathbf{B} . The last term is perpendicular to both the momentum \mathbf{p} and the magnetic induction \mathbf{B} ; hence the impulse

due to it may be thought of as a change of \mathbf{p} brought about by a rotation at the rate ω (where $\omega = d\theta/dt$) for the short time δt about the vector \mathbf{B} . Thus, as seen in Fig. 2.8, $\delta \mathbf{p}$ due to the last term can be written

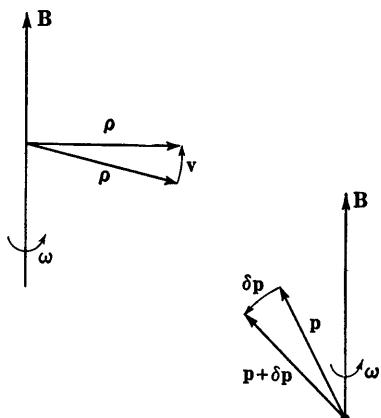


FIG. 2.8. Rotational motion of a moving charge in the presence of magnetic induction.

$$\delta \mathbf{p} = \omega \times \mathbf{p} \delta t$$

Equating this to the last term,

$$\omega \times \mathbf{p} \delta t = \frac{e}{m} \mathbf{p} \times \mathbf{B} \delta t$$

or

$$\omega = -\frac{e}{m} \mathbf{B} \quad (2.7)$$

The trajectory represented by this rotation of \mathbf{p} at the angular rate ω can be described by a radius of curvature ρ such that $\omega \times \mathbf{r} = \mathbf{v}$ or $m\omega \times \mathbf{r} = \mathbf{p}$, and hence

$$\mathbf{p} = e\mathbf{r} \times \mathbf{B} \quad \text{or} \quad \mathbf{r} = \frac{1}{e} \frac{\mathbf{B} \times \mathbf{p}}{B^2} \quad (2.8)$$

It is also interesting to see the result of integrating the scalar product \mathbf{F} and $\delta\mathbf{r}$, which yields the work done on the charge e .

$$\int \mathbf{F} \cdot \delta\mathbf{r} = e \int \mathbf{E} \cdot \delta\mathbf{r} + \frac{e}{m} \int (\mathbf{p} \times \mathbf{B}) \cdot \delta\mathbf{r}$$

Writing $\mathbf{p} = m\mathbf{v} = m\delta\mathbf{r}/\delta t$,

$$\int \mathbf{F} \cdot \delta\mathbf{r} = e \int \mathbf{E} \cdot \delta\mathbf{r} + \frac{e}{m^2} \int \mathbf{p} \times \mathbf{B} \cdot \mathbf{p} \delta t$$

The second integrand on the right vanishes showing that the constant magnetic induction \mathbf{B} contributes nothing to the changing energy. If the work done by the electric field appears as a change in kinetic energy only, then

$$\int c^2 \delta m = e \int \mathbf{E} \cdot \delta\mathbf{r} \quad (2.9)$$

Special cases of the above equations corresponding to particular geometries and field dispositions determine the trajectories of charged particles. Fields corresponding to many different choices of the potentials \mathbf{A} and φ have been studied and used to determine precise values of the charge-to-mass ratios of atomic particles. Suitable choices of the fields and geometries lead to conditions under which trajectories diverging from a source of ions may be made to converge again, yielding focusing effects which greatly enhance the precision of measurement obtainable. These focusing conditions also render operable the large, high-energy ion accelerators used for nuclear research.

The simplest cases are those in which constant, uniform electric and magnetic fields are employed. The equations then reduce to

$$p_{\perp} = e\rho B \quad \text{and} \quad (m - m_0)c^2 = e\varphi$$

or, in the nonrelativistic approximation,

$$v_{\perp} = \frac{e}{m} \rho B \quad \text{and} \quad v^2 = 2 \frac{e}{m} \varphi$$

where φ is a potential difference and v_{\perp} is the component of \mathbf{v} normal to \mathbf{B} . When the configuration is such that $v = v_{\perp}$,

$$\frac{e}{m} = \frac{2\varphi}{\rho^2 B^2} \quad (2.10)$$

Alternatively, the value of e/m may be determined by the use of either one of these equations together with some method of timing the passage of the charged particle over a known path. All the different combinations of these methods have been used in precision determinations of

the ratio e/m for electrons from many different sources. One of the most precise recent determinations is that of Dunnington,¹ who obtained 1.75982×10^{11} coulombs kg^{-1} . Indirect measurements allow additional determinations of e/m_0 , and Birge² gives $(1.75880 \pm 0.00028) \times 10^{11}$ coulombs kg^{-1} as an average of spectroscopic determinations of e/m_0 . The best indirect (least-squares adjusted) value is

$$e/m_0 = -(1.75888 \pm 0.00005) \times 10^{11} \text{ coulombs } \text{kg}^{-1}$$

Combining this with the value of e , the rest mass of the electron is found to be

$$m_0 = (9.1085 \pm 0.0006) \times 10^{-31} \text{ kg}$$

The absolute value of this universal constant of nature is of comparable significance in atomic physics to those of e and N .

Mass Spectroscopy. In early experiments by Thomson on electric conduction in gases, the *positive rays* observed to stream through an orifice in the negative electrode were analyzed by passing them through uniform regions of electric field and magnetic induction, E and B being parallel to one another and perpendicular to the ray path. With this geometry the equations for v in the preceding paragraph are seen to lead to a parabolic equation for the intersection of the surface into which an ion beam of given e/M is spread out and a plane surface normal to the beam. The constant of the parabola determines the ratio e/M . These early experiments were illuminating but not capable of great precision. Through them the existence of isotopes was first demonstrated in experiments with neon. Modifications of this geometrical arrangement to improve precision have been used by subsequent experimenters to determine the masses of the isotopes of almost all the chemical elements. Internal consistency of the experimental observations under a variation of parameters determines the number of unit charges carried by the ions, and hence the atomic mass can be determined from the charge-to-mass ratio. Since this ratio is smaller than for electrons, the value of the magnetic induction which is employed is generally larger than in experiments with electrons. Ferromagnetic materials are usually employed in the magnetic circuit, so the magnetic induction is not directly calculable and its measurement with precision presents experimental difficulties. In consequence, until recently, absolute values of the atomic masses have not been determined in this way.

However, comparative measurements can be made with a very high precision. The trajectories of two types of ions for which the masses are nearly in an integral ratio are closely the same if the charges carried

¹ F. G. Dunnington, *Phys. Rev.*, **52**, 475 (1937).

² R. T. Birge, *Repts. Progr. in Phys.*, **8**, 90 (1942).

are in the same integral ratio. Thus, if two groups of ions with charges e and ne and masses M and $n(M + \delta M)$, where δM is small, fall through the same potential difference φ and are deflected to some sharply defined focus in a region of magnetic induction B , the difference in the radii of curvature and hence of the focal points may be obtained from

$$e\varphi = \frac{e^2\rho^2B^2}{2M} \quad \text{and} \quad ne\varphi = \frac{n^2e^2(\rho + \delta\rho)^2B^2}{2(M + \delta M)n}$$

Thus to the first order of small quantities,

$$\frac{\delta M}{M} = \frac{2\delta\rho}{\rho} \quad (2.11)$$

The difference in mass δM can thus be measured with the precision of M , ρ , and $\delta\rho$ which yields a much higher over-all accuracy for relative mass values than for absolute ones. The mass of O^{16} is taken arbitrarily as 16 atomic mass units (amu). As an illustration, singly charged hydrogen can be compared with doubly charged deuterium, singly charged deuterium with doubly charged helium, and so on. The crucial mass comparisons in determining the atomic weights of the light elements are¹

$$\begin{aligned} H_2 - D &= 0.0015503 \pm 0.0000015 \text{ amu} \\ D_3^2 - \frac{1}{2}C^{12} &= 0.042292 \pm 0.000012 \text{ amu} \\ C^{12}H_4 - O^{16} &= 0.036371 \pm 0.000012 \text{ amu} \end{aligned}$$

From these and analogous doublet comparisons for other molecules, the precise atomic weights relative to $O^{16} = 16$ can be determined.² To obtain the atomic mass from that of the ion, the mass of the removed electrons must of course be added. A list of the atomic weights of the lighter permanent isotopes is given at the end of this chapter.

Nuclear Transmutations. The methods of mass measurement described in the preceding paragraphs apply only to particles with which an electric charge can be associated. One particle is known which is fundamentally uncharged, and in consequence its mass cannot be measured by such techniques. This is the *neutron*, and its existence was demonstrated by

¹ K. T. Bainbridge, *Phys. Rev.*, **81**, 146 (1951); H. Ewald, *Z. Naturforsch.*, **69**, 293 (1951). Other important recent doublet measurements are given in A. O. Nier and T. R. Roberts, *Phys. Rev.*, **81**, 507 (1951); A. O. Nier, *Phys. Rev.*, **81**, 624 (1951); T. R. Roberts, *Phys. Rev.*, **81**, 624 (1951).

² For an account of certain mass-spectrographic techniques and ion trajectories the following references are of interest: E. B. Jordan and L. B. Young, *J. Appl. Phys.*, **13**, 526 (1942); N. D. Coggeshall, *Phys. Rev.*, **70**, 270 (1946); H. G. Thode and R. B. Shields, *Repts. Progr. in Phys.*, **12**, 1 (1948); W. Walcher, *Nucleonics*, **5**, 42 (1949); W. H. Bennett, *J. Appl. Phys.*, **21**, 143 (1950); K. I. Mayne, *Repts. Progr. in Phys.*, **15**, 24 (1952); K. T. Bainbridge, pt. V, vol. I, of E. Segre, "Experimental Nuclear Physics," John Wiley & Sons, Inc., New York, 1953.

Chadwick.¹ He observed that kinetic energy was transferred by certain entities, arising from the nuclear reaction between the emanations of radium and beryllium, to the protons in a hydrogen-containing substance. A consistent description of the observed momenta acquired by the protons could be given on the assumption that they were being projected forward as a result of elastic collisions with particles of approximately their own mass. Such particles were uncharged, for they produced no ionization along their paths and were undeflected by electric or magnetic fields. A precise measurement of the neutron mass comes from the knowledge that the deuteron is composed of a neutron and proton, as shown by nuclear disintegration experiments, and from a subtraction of the proton mass from that of the deuteron plus the mass equivalent to the measured binding energy. Intercomparisons of atomic masses which are independent of mass-spectrographic measurements are afforded by other nuclear disintegration experiments. Li, Whaling, Fowler, and Lauritsen² have evaluated the precision measurements of transmutations of the light elements and have determined a consistent set of mass values. As an example of the accuracy and reliability afforded by these nuclear measurements, consider the following transmutation cycles:³

$$\begin{array}{r} \text{T}^3 + \text{H} = \text{He}^3 + n - 764 \pm 1 \text{ kev} \\ \text{T}^3 = \text{He}^3 + \beta_- + 18.6 \text{ kev} \\ \hline n - \text{H} = 782.6 \pm 1 \text{ kev} \end{array}$$

$$\begin{array}{r} \text{C}^{12} + \text{D} = \text{N}^{13} + n - 281 \pm 3 \text{ kev} \\ \text{C}^{12} + \text{D} = \text{C}^{13} + \text{H}^1 + 2725 \pm 6 \text{ kev} \\ \hline \text{N}^{13} = \text{C}^{13} + \beta_+ + 1022 + 1206 \pm 4 \text{ kev} \\ n - \text{H} = 778 \pm 4 \text{ kev} \end{array}$$

The agreement between the mass-spectrographic determination of masses through the mass-doublet measurements and nuclear transmutation data is exhibited by this cycle:

$$\begin{array}{r} {}_4\text{Be}^9 + {}_1\text{H}^1 = {}_4\text{Be}^8 + {}_1\text{D}^2 + 558 \pm 2 \text{ kev} \\ {}_4\text{Be}^8 = {}_2\text{He}^4 + {}_2\text{He}^4 + 95 \pm 1 \text{ kev} \\ {}_4\text{Be}^9 + {}_1\text{D}^2 = {}_3\text{Li}^7 + {}_2\text{He}^4 + 7150 \pm 9 \text{ kev} \\ {}_3\text{Li}^7 + {}_1\text{H}^1 = {}_2\text{He}^4 + {}_2\text{He}^4 + 17,339 \pm 12 \text{ kev} \\ \hline 2\text{D}^2 - \text{He}^4 = 23,836 \pm 10 \text{ kev} \\ = (25.599 \pm .010) \times 10^{-3} \text{ amu} \quad (\text{transmutation cycle}) \end{array}$$

¹ J. Chadwick, *Proc. Roy. Soc. (London)*, **136A**, 692 (1932).

² C. W. Li, W. Whaling, W. A. Fowler, and C. C. Lauritsen, *Phys. Rev.*, **83**, 572 (1951).

³ K. T. Bainbridge, *loc. cit.*

compared to

$$2D^2 - He^4 = (25.604 \pm .008) \times 10^{-3} \text{ amu} \quad (\text{mass doublet})$$

The present values of the light masses as given in Table 2.2 are consistent with both the nuclear data and the mass-doublet measurements of Ewald.

Resonance Methods. An accurate absolute measurement has recently been made of the ratio of the charge to the mass of a proton by Hipple.¹ The possibility of accomplishing this depends upon the ability to make a precise measurement of the resonance frequency of orbital revolution of a hydrogen ion in a region where the magnetic induction is accurately known. The value of e/M_p , where M_p is the proton mass, is subsequently calculated using Eq. (2.7). The apparatus is shown in Fig. 2.9. A small apparatus with dimensions of the order of 1 cm on a side is placed in a region of very homogeneous magnetic induction of

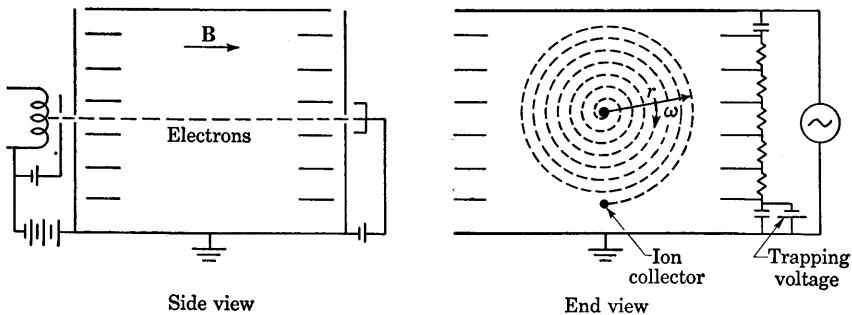


FIG. 2.9. Omegatron for measuring the cyclotron frequency. (Hipple.)

about 0.5 weber m⁻². The value of the magnetic induction is determined by making a precise absolute measurement of the force on a conductor carrying a known current. A beam of electrons is injected parallel to \mathbf{B} , and protons or hydrogen ions are produced by this beam traversing the residual gas in the region. A sinusoidal alternating electric field is produced at right angles to \mathbf{B} . This field is rendered nearly uniform over the volume by the employment of equipotential planes parallel to \mathbf{B} , the centers of which are cut out to permit the spiraling out of ions from their region of formation in the electron beam. If the relation between \mathbf{B} and the angular frequency of oscillation of the electric field is correct, the amplitude of circular motion of the ions increases until they reach a collecting electrode near the periphery of the maximum possible circular orbit and are there detected as an electric current. Assuming that the alternating electric field is in the x direction, the equations of motion are, from Eq. (1.17),

¹ J. A. Hipple, H. Sommer, and H. A. Thomas, *Phys. Rev.*, **76**, 1877 (1949).

$$\begin{aligned} M\ddot{x} &= eE_{x0}e^{i\omega t} + ejB - e\alpha\dot{x} \\ M\ddot{y} &= -eB\dot{x} - e\alpha\dot{y} \end{aligned}$$

where the applied electric field is written in complex notation¹ and α is a damping factor assumed qualitatively to account for loss of energy of the ions during their motion. If the second of these equations is multiplied by i and they are added together, $x + iy$ subsequently being written r , the resulting equation is

$$M\ddot{r} + (\alpha + iB)e\dot{r} = eE_{x0}e^{i\omega t}$$

A steady-state solution is $r_0e^{i\omega t}$ which on substitution yields for r_0

$$r_0 = \frac{E_{x0}}{\omega\{(B + (M/e)\omega)^2 + \alpha^2\}^{\frac{1}{2}}} e^{i\phi} \quad (2.12)$$

where $\phi = \tan^{-1}[\alpha/(B + \omega M/e)]$. The amplitude of rotation r_0 is evidently a maximum for the condition of Eq. (2.7), and an ion current is observed at the collector if its radius is less than $ME_{x0}/Be\alpha$. In Hippel's apparatus a peak in the curve of collector current as a function of frequency is observed with residual hydrogen ions (H_2^+) for $B = 0.4697$ weber m⁻² when $\omega = 1.126 \times 10^7$ rad sec⁻¹. The narrowness of the peak is such that a precision of about 1 part in 10^5 can be attained. Using the value of e/m_0 to correct for the additional electron, e/M_p is found to be $(9.57942 \pm 0.0003) \times 10^7$ coulombs kg⁻¹, where M_p is the proton mass. If e/m_0 for an electron is divided by this quantity, the ratio M_p/m_0 is obtained as 1836.12 which is in good agreement with the best present value for this ratio as given in the list of atomic constants. Similar agreement is obtained if this value of e/M_p is used to calculate the value of the electronic charge or the faraday.

Somewhat analogous methods of determining ionic masses by measuring characteristic angular velocities in a known magnetic induction have been proposed by Goudsmit² and developed by Goudsmit³ and his colleagues and by Smith.⁴ In the mass synchrometer of Smith and Damm, which is illustrated schematically in Fig. 2.10, ions from a gas-discharge source traverse circular paths in a region of constant magnetic induction and reach a collector after three successive passages through a short section 180° from the source in which they receive such decelerations that they pass through slits in the central baffle defining successive decreases in the radius of curvature. The ions that traverse this path

¹ The symbol e for the electron charge should not be confused with e used as the base of the natural logarithm.

² S. A. Goudsmit, *Phys. Rev.*, **74**, 622 (1948).

³ E. E. Hays, R. I. Richards, and S. A. Goudsmit, *Phys. Rev.*, **84**, 824 (1951).

⁴ L. G. Smith, *Rev. Sci. Instr.*, **22**, 115 (1951); C. C. Damm and L. G. Smith, *Phys. Rev.*, **91**, 482 (1953).

and reach the collector are those which make a complete revolution in the magnetic induction in very nearly an integral number n of cycles of the applied radio-frequency field. The ion current to the collector is amplified and applied to the vertical deflecting plates of an oscilloscope, and the frequency of the retarding signal is swept linearly through a range of values about 30 times a second. The deflecting potential applied to the horizontal oscilloscope plates is synchronized with the frequency sweep. Thus a sharp peak corresponding to the reception of ions by the collector is registered on the oscilloscope screen for each value of n and for each e/M corresponding to the species of ions emerging from the source. By making a slight displacement of the frequency range between alternate sweeps, two e/M peaks can be brought into precise coincidence.

The accuracy of peak matching and frequency measurement is such that for values of n of the order of 150 the difference in mass between closely spaced ion peaks such as N_2 and CO can be determined with a precision of about 1 part in 10^5 .

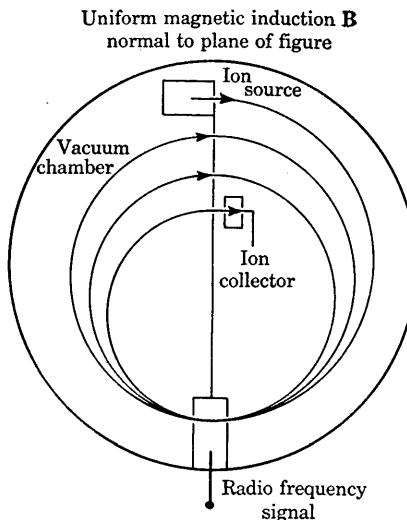


FIG. 2.10. Mass synchronometer. (*Smith.*)

2.4. Corpuscular Nature of Radiation. Radiation is essentially the term which is used to describe the mechanism through which a change in one charge configuration induces a change in another charge distribution. Momentum, angular momentum, and energy can all be interchanged between charged particles, and it is a useful and simplifying concept to assume that these physical quantities are stored in the radiation field between processes that are described as emission of radiation by one charge system and absorption of it by another. The classical analysis of electromagnetic fields to describe the phenomena of physical optics represented a triumph for the wave concept of radiation. Further, the basic implication of continuity involved in the field description is, on the surface, inconsistent with any discrete structure. In fact, phenomena such as optical interference became a test for wave transmission. Yet it was in the thermal radiation spectrum that the first suggestion of quantization originated. Classical equipartition of energy applied to thermal radiation in temperature equilibrium with its enclosing walls predicted a spectrum which tended to infinite intensity in the ultra-

violet or high-frequency end and thus contradicted experiment. In order to limit theoretically the energy drawn into this "ultraviolet catastrophe," Planck¹ in 1900 suggested a quantization of the postulated oscillators with which the radiation was assumed to be in equilibrium. Observations on the liberation of electrons from a metal surface under the influence of light (photoelectric effect) were interpreted by Einstein² in 1905 as involving the quantized absorption of light. The emission and absorption of quanta of energy in the form of radiation were the essential parts of Bohr's atom model in 1913. These were early steps in the development of the quantum or photon hypothesis of radiation, which may be stated as: The interaction of charges through the medium of radiation is accomplished by the emission and absorption of discrete quanta of radiant energy of magnitude $h\nu$, where ν is the frequency of the radiation and h is a universal constant called Planck's constant.

In order that this hypothesis may be in accord with electromagnetic theory, it is clearly necessary that the photon trajectories shall be guided by the field theory in the sense that Poynting's vector shall determine at least statistically the number of photons traversing unit area per unit time normal to the beam and that all photons shall travel in free space with the velocity of light c . A consequence of the latter fact is that if $h\nu$ is the energy of a photon its momentum is $h\nu/c$, or h/λ . If the photon is thought of as a particle with a velocity c , it is clear from Eq. (1.16) that it has no rest mass; in the case of a photon, m_0 equals zero. The photon is essentially unique among entities to which the particle concept is applied in that there is no nonrelativistic approximation by which its properties may be described. It is also of a transitory nature; it is emitted and absorbed by charged particles having finite rest masses.

The earliest direct experiment for the quantitative determination of the constant of proportionality between photon energy and frequency was that of Millikan³ on the emission of electrons from metallic surfaces under the influence of light. In essence it consisted in determining the kinetic energy of liberated electrons as a function of the frequency of the light striking the metal surface by measuring the electric potential which would just suffice to prevent the electrons from reaching a second metallic plate in the evacuated tube in which the experiment was performed. Such an experiment, illustrated in Fig. 2.11, is somewhat complicated by the necessity of allowing for the energy that it is found must be supplied to enable an electron to leave the plate against the retaining forces such as that of the image charge. This correction energy is a function of the nature of the surface, and, as surface conditions are very difficult to con-

¹ M. Planck, *Verhandl. deut. physik. Ges.*, **2**, 237 (1900); *Ann. Physik*, **4**, 553 (1901).

² A. Einstein, *Ann. Physik*, **17**, 132 (1905).

³ R. A. Millikan, *Phys. Rev.*, **7**, 355 (1916).

trol and reproduce, it constitutes a major source of uncertainty in the experiment. However, assuming that the energy delivered by a photon to an electron is $h\nu$, that the energy required for the electron to leave the neighborhood of surface A is $e\phi_A$, and finally that V is the potential difference between plates A and B which just suffices to prevent the flow of electrons, the energy equation becomes

$$h\nu = e\phi_A + eV$$

From the sketch in Fig. 2.12 of the energy relations in the circuit,

$$\varphi_A + V = \varphi_B + V'$$

where V' is the potential difference between the plates as measured by a voltmeter. Thus

$$\frac{h}{e}\nu = \varphi_A + V = \varphi_B + V' \quad (2.13)$$

The ratio h/e can be determined from the slope of the linear relation found experimentally to exist between ν and V' . φ_A and φ_B can also be determined from the data. The linearity confirms Planck's hypothesis, and the value of h/e yields a value of h which is accurate to about 1 per cent.

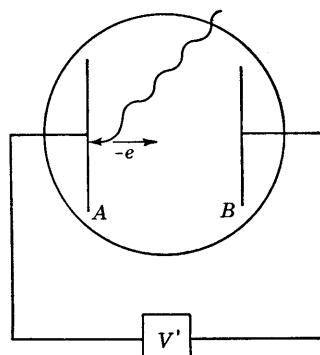


FIG. 2.11. Schematic of photoelectric-effect experiment.

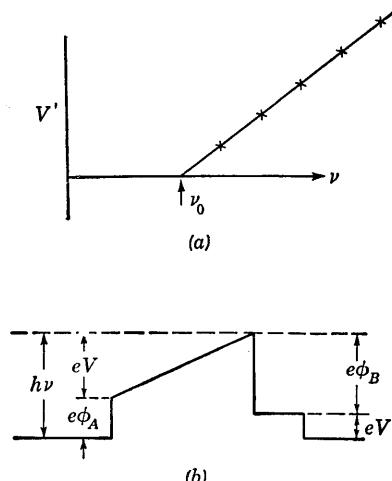


FIG. 2.12. (a) Stopping potential as a function of ν . (b) Energy relations in the photoelectric circuit.

An effect which yields further confirmation of the general correctness of the photon concept is that which is observed during the interaction of a high-energy X-ray photon with an electron that is comparatively free of restraining atomic forces. This is known as the Compton¹

¹ A. H. Compton, *Bull. Natl. Research Council (U.S.)* No. 20, p. 16 (1922); *Phys. Rev.*, **21**, 715 (1923); *Phys. Rev.*, **22**, 409 (1923).

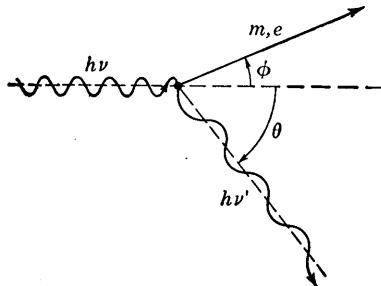
effect and can be observed and studied in Wilson-chamber photographs and by means of X-ray spectrometers. It is not a precision method for determining h , but it displays very clearly the application of particle considerations to photons in their interaction with electrons. If ν is the photon frequency, the equations for the conservation of momentum and energy with an electron, assumed to

be initially at rest, are

$$\frac{h\nu}{c} = \frac{h\nu'}{c} \cos \theta + mv \cos \phi$$

$$0 = \frac{h\nu'}{c} \sin \theta - mv \sin \phi$$

$$h\nu = (m - m_0)c^2 + h\nu'$$



where ν' is written for the possibly altered frequency of the scattered

FIG. 2.13. Compton effect.

photon and the geometry is depicted in Fig. 2.13. Eliminating ϕ , an equation is obtained between ν , ν' , θ and the constants. This may be written more conveniently in terms of the wavelengths $\lambda = c/\nu$ and $\lambda' = c/\nu'$ as

$$\lambda' - \lambda = \frac{h}{m_0 c} (1 - \cos \theta) = \frac{2h}{m_0 c} \sin^2 \frac{\theta}{2} \quad (2.14)$$

This equation explains the assumption of a changed frequency λ' , for if $\lambda' = \lambda$ there is no deflection or other change in the radiation. The greatest value of $\lambda' - \lambda$ is $2h/m_0 c$, where θ has the value π ; hence

$$\frac{\nu - \nu'}{\nu} = \left(1 + \frac{m_0 c^2}{2h\nu} \right)^{-1}$$

is the maximum fractional energy the photon can lose at such a collision. Thus it cannot be completely absorbed, and it is necessary to postulate the existence of a scattered photon of energy $h\nu'$. The quantity $h/m_0 c$, which is the change in wavelength for 90° scattering, is known as the *Compton wavelength*. It is seen to be the wavelength of radiation having a quantum energy equivalent to the rest mass of an electron, and it is greater than the classical radius r_0 of Sec. 1.8 by the dimensionless factor $4\pi\epsilon_0 ch/e^2 = 2\pi/\alpha$ where α , which is known as the *fine-structure constant*, is about $1/137$. Equation (2.14) is well confirmed by experiment, and this agreement is a most cogent argument for the corpuscular nature of radiation and the application of the laws of conservation of energy and momentum to photon processes. This is, of course, not the only process that takes place when X rays are scattered by matter. Essentially elastic scattering (no energy loss) takes place by tightly bound electrons

for which the mass in the denominator of Eq. (2.14) is essentially the atomic mass and hence $\lambda' - \lambda$ is very small. Also, characteristic energy losses occur which are associated with atomic energy levels.

The most precise determinations of the value of h come from indirect measurements such as determinations of the Rydberg constant (Sec. 3.1) and from high-energy quantum and diffraction processes. The measurements of the magnetic deflection of photoelectrons ejected from metals by characteristic X rays yield a precise value for $\frac{e}{h} \frac{e}{m_0}$. Robinson¹ and his colleagues have used the silver $K_{\alpha 1}$ and other X-ray lines to eject electrons from the K , L , or M levels of gold, platinum, silver, or copper. These photoelectrons were deflected in a beta-ray spectrograph, and the $H\rho$ values determined. The total energy of the electrons is $E = h\nu - W + m_0c^2 = hc/\lambda - W + m_0c^2$, and their momentum is $p = Be\rho$, where the X-ray wavelengths λ are well known from diffraction measurements, the $B\rho$ values are measured, and W is the binding energy of the electron. If these relations are inserted in Eq. (1.29),

$$\begin{aligned}\frac{hc}{\lambda} - W + m_0c^2 &= [(B\rho e c)^2 + (m_0c^2)^2]^{\frac{1}{2}} \\ \frac{1}{\lambda} - \frac{W}{ch} &\cong \frac{e^2}{hm_0} \cdot \frac{B^2\rho^2}{2c}\end{aligned}$$

The binding energy W can be eliminated by using several X-ray lines.

Measurements of electron diffraction by crystals as discussed in Sec. 2.5 also involve h . The wavelength λ of these waves can also be determined by the diffraction patterns produced by the crystal lattice spacings. Using Eqs. (1.29) and (2.16),

$$\begin{aligned}(e\varphi + m_0c^2)^2 &= p^2c^2 + m_0^2c^4 \\ e\varphi &\cong \frac{h^2}{2m_0\lambda^2}\end{aligned}$$

or

$$2\varphi\lambda^2 \cong \frac{h^2}{em_0}$$

where φ is the measured acceleration voltage.

The most precise measurements involving h have been made of the short-wavelength limit of the continuous X-ray spectrum.² A recent experiment³ used a sealed-off tungsten X-ray tube operated at 24,498.7

¹ C. J. B. Clews and H. R. Robinson, *Proc. Roy. Soc. (London)*, **176**, 28 (1940).

² W. K. H. Panofsky, A. E. S. Green, and J. W. M. DuMond, *Phys. Rev.*, **62**, 214 (1942); J. A. Bearden, F. T. Johnson, and H. M. Watts, *Phys. Rev.*, **81**, 70 (1951).

³ G. L. Felt, J. N. Harris, and J. W. M. DuMond, *Phys. Rev.*, **92**, 1160 (1953).

± 1.1 volts (the cathode work function of 4.52 volts must be added to this). The wavelength of the quantum limit of the X rays was measured with DuMond's 2-m focusing curved-quartz-crystal spectrometer, by comparison with the tin $K_{\alpha 1,2}$ and silver $K_{\beta 1,2}$ lines. A least-squares adjusted value of h consistent with such experiments and the other atomic constants has been calculated by DuMond and Cohen¹ as

$$h = (6.6252 \pm 0.0005) \times 10^{-34} \text{ joule-sec}$$

The value of this important universal constant has until recently been somewhat less precisely known than were those of N or e because its accurate determination has presented greater experimental difficulties.

2.5. The Wave Nature of Matter. The conceptual parallelism between photons and material atomic particles was greatly enhanced by the interpretation of experiments made during the latter part of the 1920's on the scattering of elementary particles by crystal lattices. The first experiments were those of Davisson and Germer² in 1927 on the scattering of a narrow beam of electrons incident normally on the (111) plane of a nickel crystal (see Fig. 2.14). The scattered electrons observed were those which had lost no energy, analogous to coherent X-ray scattering. When the nickel surface was as clean and free of impurity layers as possible, sharp maxima of scattering were observed at certain polar angles with the beam and at azimuthal angles displaying the sixfold symmetry characteristic of the (111) plane of nickel. The electron energies were low, of the order of 50 to 150 volts, so they were unable to penetrate far below the crystal surface and still be scattered out again. Nevertheless, there was evidence of the enhancement of certain surface-grating maxima by scattering from deeper atomic layers. For any particular electron energy the scattering maxima observed were in accord with the scattering of a wave, comparable in length with the lattice spacing, which had a somewhat different length in the crystal than outside. On changing the electron energy to a different value the apparent wavelength changed in

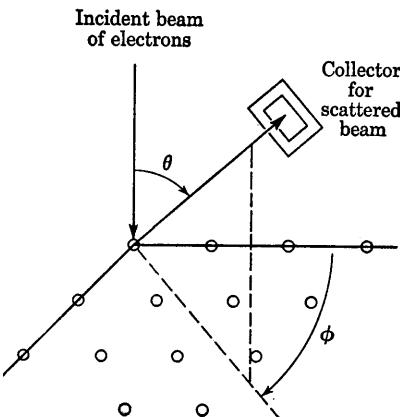


Fig. 2.14. Exploration of the diffraction pattern of electrons from a nickel crystal.

¹ J. W. M. DuMond and E. R. Cohen, *Rev. Mod. Phys.*, **25**, 691 (1953).

² C. J. Davisson and L. H. Germer, *Phys. Rev.*, **30**, 705 (1927); *Proc. Natl. Acad. Sci. U.S.*, **14**, 317, 619 (1927).

such a way as to be inversely proportional to the square root of the energy, i.e., inversely proportional to the momentum. Finally, the product of the wavelength and momentum was of the order of Planck's constant.

This was the first recognized experimental evidence of the wave nature of a material particle, but as the electron energies were not very much greater than the energy of binding between an electron and crystal the quantitative results were not very precise. The use of much higher electron energies by G. P. Thomson,¹ Rupp,² and Kikuchi³ verified with considerable precision that the product of the electron momentum and wavelength is equal to \hbar . Diffraction of electrons by ruled gratings was also observed,⁴ and crystal diffraction experiments by von Friesen⁵ have yielded a precision value of \hbar/\sqrt{em} . The techniques of such a measurement are so similar to X-ray diffraction techniques that they need not be described in detail. The theory is the simple one of coherent scattering from lattice centers. The patterns of scattering maxima are clearly distinguishable from X-ray patterns in that they are deflected by an electric field and a magnetic induction with the characteristic parameter e/m_0 .

Wave properties of material particles with finite rest masses are not limited to electrons. Experiments by Estermann and Stern⁶ showed that hydrogen molecules and helium atoms when scattered from a crystal surface display characteristic diffraction maxima. Johnson⁷ has shown this to be true for atomic hydrogen, and Zinn⁸ and his colleagues for neutrons. Thus it appears reasonably certain that the momentum-wavelength relation is a fundamental one, characteristic of all material particles. In consequence, the following equations may be taken as applying to all types of particles, photons, or material, with either a zero or positive rest mass.

$$E = mc^2 = \hbar\nu = \hbar\omega \quad (2.15)$$

$$p = mv = \frac{\hbar}{\lambda} = \hbar k \quad (2.16)$$

where $\omega = 2\pi\nu$, $k = 2\pi/\lambda$, and $\hbar = h/2\pi$. Since the product of ν and λ

¹ G. P. Thomson, *Proc. Roy. Soc. (London)*, **117**, 600 (1928); **119**, 651 (1928).

² E. Rupp, *Ann. Physik*, **85**, 981 (1928).

³ S. Kikuchi, *Physik Z.*, **31**, 777 (1930).

⁴ B. L. Worsnop, *Nature*, **123**, 164 (1929); E. Rupp, *Z. Physik*, **52**, 8 (1928).

⁵ S. von Friesen, Uppsala Dissertation (1936); *Proc. Roy. Soc. (London)*, **160**, 424 (1937).

⁶ I. Estermann and O. Stern, *Z. Physik*, **61**, 95 (1930).

⁷ T. H. Johnson, *J. Franklin Inst.*, **206**, 301 (1928); **207**, 639 (1929).

⁸ W. H. Zinn, *Phys. Rev.*, **70**, 102 (1946); E. O. Wollan, E. G. Shull, and M. C. Marney, *Phys. Rev.*, **73**, 527 (1948).

is by definition the phase velocity v_p of a wave, $E/p = v_p = c^2/v$. Thus the product of the particle velocity v and the phase velocity of the associated wave is equal to the square of the velocity of light. Since v is less than c , v_p is greater than c , but this is not unacceptable since information is not propagated with the velocity v_p .

The wave properties of material particles can be taken as well established by diffraction experiments, but these provide little help in formulating a concept of what pulsates and confers this periodic property on particles. Some insight is provided by the speculations of de Broglie,¹ who in 1924 considered the possibility of associating a characteristic pulsation with a mass. This was prior to the experimental observation of matter waves. Without inquiring too closely into any possible structural attributes responsible for the wave nature of a particle, assume that an observer at rest with respect to a particle would describe some characteristic vibration taking place there in terms of his proper time by the function ψ , where $\psi = \psi_0 e^{2\pi i \nu_0 t_0}$ and ν_0 is the frequency he observes. ψ may be thought of as a function describing the physical state of the system under consideration, in this case a free particle. An observer moving with respect to the particle with a velocity v would, in consequence of the time transformation [Eq. (1.13)] $t_0 = (t - vz/c^2)/(1 - v^2/c^2)^{1/2}$, describe the function ψ as

$$\psi = \psi_0 \exp \left[\frac{2\pi i \nu_0 (t - vz/c^2)}{\sqrt{1 - v^2/c^2}} \right]$$

Thus he would describe his observation as a wave having the frequency $\nu = \nu_0(1 - v^2/c^2)^{-1/2}$ and traveling with a phase velocity $(v/c^2)^{-1}$. This is closely analogous to the situation as disclosed by the experiments on the diffraction of material particles, and it is suggestive, since the transformation of ν and m is the same, that a natural frequency is associated with the rest mass of a particle. From Eqs. (2.15) and (2.16) the relation between m and ν is seen to be $\nu_0 = m_0 c^2/h$. The equation for ψ may be generalized in terms of the radius vector \mathbf{r} and propagation vector \mathbf{k} or in terms of the energy and momentum as

$$\psi = \psi_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})} = \psi_0 e^{-i(Et - \mathbf{p} \cdot \mathbf{r})/\hbar} \quad (2.17)$$

The above discussion is of course merely suggestive as it is limited to the description of the simplest situation imaginable: one particle in the universe moving in the direction defined by \mathbf{p} but without any spatial localization, since the domain of ψ is unrestricted.

A certain degree of spatial localization, which is desirable if the wave pattern is to represent in some sense a particle, can be brought about by

¹ L. de Broglie, *Phil. Mag.*, **47**, 446 (1924).

postulating many waves which interfere constructively with one another in a spatial region. The nature of the results depends little on the precise choice of distribution of amplitude with wave number among these waves, and for mathematical simplicity it will be assumed that the amplitudes are distributed in accordance with the error function $\psi_{00} e^{-(k-k_0)^2/2a^2} dk$, where a is a parameter. The amplitude of the wave at $k = k_0 \pm \sqrt{2}a$ is $1/e$ of the amplitude for $k = k_0$, as shown in Fig. 2.15. Choosing z as the axis

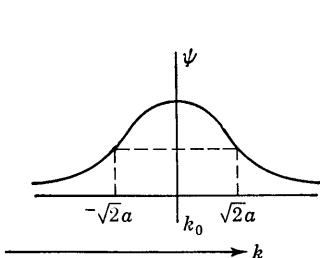


FIG. 2.15. Variation of wave amplitude with the wave number k .

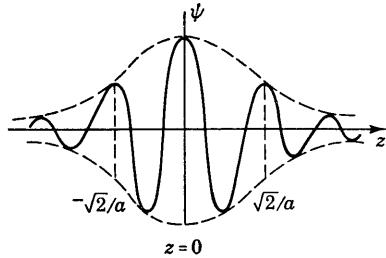


FIG. 2.16. Variation of wave amplitude with the coordinate z .

of propagation and fixing attention at $t = 0$, the resultant ψ is the sum of all waves from $k = -\infty$ to $k = \infty$, or

$$\begin{aligned}\psi &= \psi_{00} \int_{-\infty}^{\infty} e^{-(k-k_0)^2/2a^2} e^{ikz} dk \\ &= \psi_{00} \int_{-\infty}^{\infty} e^{ik_0 z - z^2 a^2/2} e^{-\alpha^2/2a^2} d\alpha\end{aligned}$$

where $\alpha = (k - k_0) - ia^2 z$. The first term is independent of α , and, since $\int_{-\infty}^{\infty} e^{-\alpha^2} d\alpha = \sqrt{\pi}$, then

$$\psi = \psi_{00} \sqrt{2\pi} a e^{ik_0 z} e^{-z^2 a^2/2}$$

Figure 2.16 exhibits ψ as a function of z and shows that the amplitude of the wave $e^{ik_0 z}$ varies with z as $e^{-z^2 a^2/2}$. Hence spatial localization is indeed achieved in that the amplitude of ψ at $z = \pm \sqrt{2}/a$ is $1/e$ of its value at $z = 0$. The product of the $1/e$ half widths of the k distribution and the z distribution is seen to be a constant independent of a . The constant a is a parameter which is a direct measure of the width in k and narrowness in z or vice versa. Additional localization in the other two dimensions, x and y , to form a three-dimensional packet of waves can be obtained by assuming that additional waves are added with their propagation vectors \mathbf{k} at slight angles with the z axis, so that $\mathbf{k} \cdot \mathbf{r}$ of Eq. (2.17) becomes $k_x x + k_y y + k_z z$, and by assuming error-function distributions of k_x and k_y about zero. The same conclusions are then reached about localization in the x and y direction as were found above for the z axis.

The interpretation of ψ , which was introduced by Born,¹ is that the square of the absolute magnitude of ψ is a measure of the probability of observing the particle in a given range of the spatial variables, or within a given range of the components of the propagation vector. This is analogous to the interpretation of the square of the field strength, or Poynting's vector, as a measure of the number of photons passing through unit area per unit time normal to a beam. The justification of this assumption by Born is of course to be looked for in the agreement between the resulting theory and experimental observation. The parallel between particle and photon diffraction patterns and the success of the theory in describing other atomic phenomena provide this justification. Hence the root-mean-square values of the widths of the ψ maxima in Figs. 2.15 and 2.16 as functions of the coordinates and as functions of the propagation vector or momentum components are of particular interest. For these are then measures of the probability of observing the particle within the stipulated ranges of the coordinates or momenta. Since the distribution of ψ in the momentum coordinate $k' = k - k_0$ is given by $Ae^{-k'^2/2a^2}$, where A is arbitrary, the mean-square value of k' is

$$\begin{aligned}\overline{k_z'^2} &= \frac{A \int_{-\infty}^{\infty} k_z'^2 e^{-k_z'^2/2a^2} dk_z'}{A \int_{-\infty}^{\infty} e^{-k_z'^2/2a^2} dk_z'} \\ &= \frac{(2a^2)^{\frac{3}{2}} \int \alpha^2 e^{-\alpha^2} d\alpha}{(2a^2)^{\frac{1}{2}} \int e^{-\alpha^2} d\alpha}\end{aligned}$$

Integrating the numerator by parts, it is seen that

$$\sqrt{\overline{k_z'^2}} = \Delta k_z' = \Delta k_z = a$$

Similarly, using the distribution in z and determining $\sqrt{\overline{z^2}}$, which is written Δz , it is found that Δz is equal to $1/a$. Therefore $\Delta k_z \Delta z = 1$. The generalization of this result in terms of the three coordinates and momenta yields what are known as the conditions of uncertainty and complementarity for a wave representing a particle.

$$\begin{aligned}\Delta x \Delta p_x &= \hbar \\ \Delta y \Delta p_y &= \hbar \\ \Delta z \Delta p_z &= \hbar\end{aligned}\tag{2.18}$$

On applying these equations to a particle moving in a circular arc, it is clear that they also imply that the product of the uncertainty in angular location of a particle and its uncertainty in angular momentum is equal to \hbar (see Fig. 2.17).

In view of the arbitrariness of the particular variation of ψ with k

¹ M. Born, *Z. Physik*, **37**, 863; **38**, 803 (1926).

assumed to achieve localization, the actual equalities are not significant. Different choices would result in the appearance of a factor slightly greater than unity on the right. The conclusion, due to Heisenberg,¹ that the product of the uncertainty in position and the uncertainty in the conjugate momentum is of the order of \hbar is a general result of basic significance in all physical phenomena. These two quantities cannot be determined separately with arbitrary accuracy, for precision in one implies a lack of it in the other. It should also be noted that an analogous argument relating to temporal rather than spatial localization implies a wave packet made up of a range of angular frequencies, and the following expression is found to hold:

$$\Delta\omega \Delta t = 1 \quad \text{or} \quad \Delta E \Delta t = \hbar \quad (2.19)$$

This states that a localization of a particle in a temporal interval implies an uncertainty in the energy it possesses. This is immediately obvious

when the method of measuring the frequency of a vibrating system is considered. To determine the frequency precisely, a large number of complete periods must be permitted to elapse. The uncertainty in the measurement becomes less as the time available for the measurement increases.

One other feature of the wave-packet representation of a particle is important for assuring that the motion of the packet corresponds to the motion of the particle, i.e., that the rate of change of position of the center of the packet is the same as the particle velocity. The packet will of

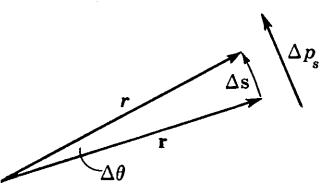
FIG. 2.17. Angle-angular momentum complementary variables.

course spread out as time progresses because of the uncertainty in momentum or velocity, but the velocity of the maximum should correspond to the particle velocity. The packet moves with the group velocity² $v_g = \partial\omega/\partial k$, and, from Eqs. (2.15) and (2.16) and the relativistic energy equation $(E/c)^2 - p^2 = (m_0c)^2$, v_g is seen to be kc^2/ω or c^2/v_p . This is the same as the particle velocity v of Eq. (2.16). Hence the particle and packet velocities are the same. A further point worth noting is that, since only $|\psi|^2$ is endowed with physical meaning in the sense of being experimentally observable, the phase of ψ is arbitrary.

2.6. Angular Momenta and Magnetic Moments of Particles. The preceding sections have been concerned with what may be considered the

¹ W. Heisenberg, *Z. Physik*, **43**, 172 (1927).

² G. Joos, "Theoretical Physics," p. 58, G. E. Stechert & Company, New York, 1934.



$$\Delta L = r \Delta p_s = \frac{r\hbar}{\Delta s} = \frac{\hbar}{\Delta\theta}$$

$$\text{or } \Delta L \Delta\theta = \hbar$$

most basic properties of atomic particles in the sense that these properties are essential to a description of such particles and to an extension of this description to the additional properties discussed in the present section. Elementary classical mechanics and electricity deal with point masses and point charges, and these are the atomic attributes that were described in Secs. 2.2 and 2.3. There are other important atomic characteristics which are essentially higher-order mechanical and electric moments, which are quite analogous to the extensions of mechanics and electricity to include the motion of rigid charged bodies. Certain general considerations of these higher-order moments will be presented in a following section. The simplest extension beyond the concept of point masses and charges, which is of primary importance for the description of the properties of such elementary atomic particles as electrons, protons, and neutrons, is that which is familiar in the classical rotation of rigid charged bodies.

Consider first the classical characteristics of the motion of a body having a certain fixed distribution of charge and mass throughout its volume. There is an intimate relation between the magnetic moment and angular momentum of such a system. The definition of magnetic dipole moment given in Sec. 1.7 can be extended to include the motion of an arbitrary charge distribution. In general, the magnetic moment is the sum of the magnetic moments for each of the charges e_i making up the total charge, or the integral of the magnetic moment associated with a continuous charge density ρ_e over the volume of the body. Thus

$$\mathbf{u} = \frac{1}{2} \sum_i e_i \mathbf{r}_i \times \mathbf{v}_i = \frac{1}{2} \int \rho_e \mathbf{r}_i \times \mathbf{v}_i dV$$

The angular momentum of the body is the sum of the angular momenta of the component masses m_i or, alternatively, the integral over the mass density. Thus

$$\mathbf{l} = \sum_i m_i \mathbf{r}_i \times \mathbf{v}_i = \int \rho_m \mathbf{r}_i \times \mathbf{v}_i dV$$

If the ratio e_i/m_i or ρ_e/ρ_m is the same for every constituent particle or for every point throughout the body, then it is seen that

$$\frac{\mathbf{u}}{\mathbf{l}} = \gamma_l = \frac{e}{2m} \quad (2.20)$$

The scalar constant γ_l is equal to the uniform ratio $e_i/2m_i$, or $\rho_e/2\rho_m$, throughout the body, and e and m may be taken as the total charge and total mass, respectively. Within the limitation of the assumed equality of the ratio for every component part of the body under consideration, Eq. (2.20) is a very general electromechanical relationship for any

type of motion. In particular, it applies to the simple type of rotational motion about an axis. The Bohr theory of atomic structure, which was based on the classical concept of circulating electrons obeying such a relationship, provided many tantalizing partial agreements with experiment but lacked several essential concepts. The concept of magnetic moment has, however, remained of primary importance in the evolution of atomic theory, and it has been found that atomic constituents have associated with them electric, magnetic, and mechanical moments which are as distinguishing as charge and mass. The static, centrally symmetric charge and mass of earlier sections may be thought of as the zero-order moments of these quantities; rotations of the charge and mass and distributions of these quantities of lesser symmetry lead to higher-order moments. The distinguishing feature of mechanical and magnetic moments in atomic theory is that they assume a unique and fundamental importance, each in its own right. The classical relation [Eq. (2.20)] between them is found to hold in some instances but not in all.

The classical nonrelativistic motion of a rigid rotating charged mass in the presence of a magnetic induction \mathbf{B} may be derived from Eq. (1.29). Writing H for the energy function in terms of the coordinates and momenta, and expanding in powers of $(v/c)^2$, the first terms are seen to be

$$H \cong m_0 c^2 + \frac{(\mathbf{p} - e\mathbf{A})^2}{2m_0} + e\varphi$$

The first of these is a constant term and may be omitted as contributing nothing to the dependence of H on the coordinates or momenta, and for the present purpose it may be assumed that there are no additional charges present to give rise to a potential, so that φ vanishes. If it is further assumed that the magnetic induction \mathbf{B} is constant in time and uniform throughout the spatial volume under consideration, \mathbf{A} may be written as $\frac{1}{2}\mathbf{B} \times \mathbf{r}$, as may be seen by forming $\mathbf{B} = \nabla \times \mathbf{A}$ with this value. Thus, omitting the subscript of m since the nonrelativistic case only is being considered,

$$\begin{aligned} H &= \frac{1}{2m} (\mathbf{p} - \frac{1}{2}e\mathbf{B} \times \mathbf{r})^2 \\ &= \frac{\mathbf{p}^2}{2m} - \frac{1}{2m} e\mathbf{p} \cdot \mathbf{B} \times \mathbf{r} + \frac{1}{8} \frac{e^2}{m} (\mathbf{B} \times \mathbf{r})^2 \end{aligned} \quad (2.21)$$

In most cases of interest these terms decrease rapidly in magnitude in the order in which they are written. This may be seen by considering the ratio eBr/p . Using the value of the electronic charge for e ; 1 weber m^{-2} for B , which is a rather large practical value, and taking r of the order of atomic dimensions, say 10^{-10} m , the numerator is of the order of

1.6×10^{-29} . If the motion is confined to the dimensions of 10^{-10} m, Eq. (2.18) indicates that p is of the order of $\hbar \times 10^{10}$, or 10^{-24} , and the quotient is of the order of 10^{-5} . Thus the terms become successively smaller as stated, and unless there is no net momentum the third term may be disregarded to a first approximation. Since $\mathbf{p} \cdot \mathbf{B} \times \mathbf{r} = \mathbf{B} \cdot \mathbf{r} \times \mathbf{p}$ and $\mathbf{r} \times \mathbf{p} = 1$, where 1 is the angular momentum, the equation may be written approximately as

$$H = \frac{1}{2m} p^2 - \frac{e}{2m} \mathbf{B} \cdot \mathbf{l} \quad (2.21')$$

Using Eq. (2.20), the second term may be written as $-\mathbf{B} \cdot \mathbf{u}$; thus $-\partial H / \partial \mathbf{B} = \mu$. Indeed, this relation may be seen to be more generally true by using the definition of μ from Sec. 1.7 and differentiating Eq. (2.21) with respect to B .

It is somewhat more informative to compare Eq. (2.21) or (2.21') with the simple free-particle Hamiltonian expressed in terms of slowly rotating coordinates. If a set of coordinates is rotating with the angular velocity ω' with respect to a primed set of coordinates, the velocities are related by the following expression, as may be seen from Fig. 2.18:

$$\mathbf{v}' = \mathbf{v} + \omega' \times \mathbf{r}$$

Thus, since $\mathbf{p} = m\mathbf{v}$,

$$\begin{aligned} H &= \frac{p'^2}{2m} = \frac{1}{2m} (\mathbf{p} + m\omega' \times \mathbf{r})^2 \cong \frac{p^2}{2m} + \mathbf{p} \cdot \omega' \times \mathbf{r} \\ &= \frac{p^2}{2m} + \omega' \cdot \mathbf{l} \end{aligned}$$

Comparing this with Eq. (2.21) or (2.21'), it is evident that the motion of a classical system of charged masses in the presence of a magnetic induction \mathbf{B} is the same, to this approximation, as if the system in the absence of \mathbf{B} were rotating as a whole with the angular velocity $\omega' = -e\mathbf{B}/2m$. By comparing this with Eq. (2.20), it is seen that $\omega'/B = -\mathbf{u}/l$. In the classical case considered here this ratio has the value of $-e/2m$. The validity of the identification of the two Hamiltonian increments $-\mathbf{u} \cdot \mathbf{B}$ and $\omega' \cdot \mathbf{l}$, however, is more general than simply for the classical case to which Eq. (2.20) applies. In a rigid rotating system of charged masses it would be anticipated that \mathbf{u} and \mathbf{l} would be

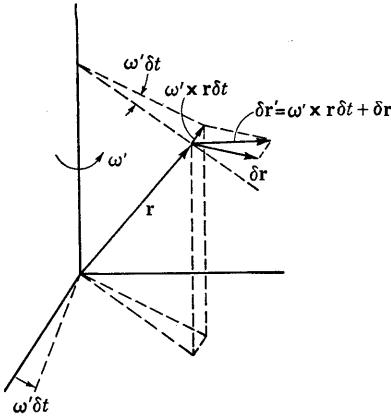


Fig. 2.18. Relation between vector increments in a stationary and rotating coordinate system.

collinear so that their ratio would be a scalar, say γ . In consequence one would have

$$\omega' = -\gamma \mathbf{B} \quad (2.22)$$

Thus, in general, a simple, rigid rotating charged mass would precess at the rate ω' in the presence of a magnetic induction \mathbf{B} , as in Fig. 2.19. The constant γ would be the same as the ratio \mathbf{u}/l , which for a classical system is $e/2m$. This simple general relationship is very important and will be extensively used in the following discussion. The classical precessional frequency ω' is known as the *Larmor frequency*. Comparing this

value of ω' with ω of Eq. (2.7), which is known as the *cyclotron frequency*, it is seen that $\omega = 2\omega'$. If the system under consideration is not spherically symmetrical, the precessional motion will be accompanied by a nutation as in the case of a spinning top in a gravitational field.¹

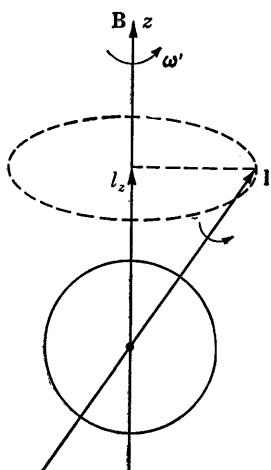


FIG. 2.19. Precession of a rotating electromechanical system in the presence of a magnetic induction.

axis, there will on the average be a finite component of \mathbf{u} along the z axis only. Deriving the force from $\mathbf{F} = -\nabla(-\mathbf{u} \cdot \mathbf{B})$ and recalling that \mathbf{B} is along the axis of z and that \mathbf{u} is not a function of the coordinates, $\mathbf{F} = F_z = \mu_B \partial B / \partial z$. The quantity μ_B is the component of \mathbf{u} along the direction of the magnetic induction, that is, $\mu_B = \mu \cos \theta$ as in Fig. 2.20.

Electrons. On the basis of these preliminaries concerning the motion of a classical material particle possessed of angular and magnetic moments, the properties of elementary particles with such moments can be discussed. In the first place it is found that all charged particles exhibit the cyclotron frequency when pursuing orbits in a region of magnetic induction. On a large scale this forms the basis of the ion cyclotrons in which charged particles are accelerated to high energies by subjecting them to periodic accelerations with the frequency ω . One type of intra-

¹ H. Goldstein, *Am. J. Phys.*, **19**, 100 (1951).

atomic electronic motion which may be thought of in terms of electrons pursuing orbits about atomic nuclei also exhibits the Larmor frequency of precession in the presence of magnetic induction. As mentioned earlier, however, these simple concepts were found inadequate to account for all the phenomena of atomic spectra. Uhlenbeck and Goudsmit¹ in 1925 postulated that the electron is possessed of an inherent spinning motion for which the ratio of the precession frequency ω' to the magnetic induction \mathbf{B} is about twice the classical value. This was a very successful and fruitful hypothesis which went far toward the elucidation of atomic spectra and also endowed the electron with a very characteristic and important though essentially non-classical attribute. The present quantum theory of an electron in a hydrogen atom, outlined in Sec. 4.3, provides more general theoretical support for this hypothesis by a suitable relativistic formulation of the problem.

The evidence for associating an intrinsic angular momentum and magnetic moment with an electron, in addition to any angular momentum or magnetic moment arising from an orbital type of motion, comes largely from the study of electronic behavior in atomic systems. Because of the magnitude of the amperian force upon the charge of an electron moving through a region of magnetic induction, it is infeasible to measure the force

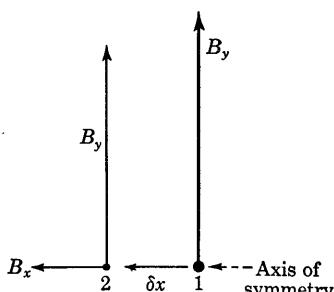


FIG. 2.21. Existence of the component B_x off the axis of symmetry necessitated by Poisson's equation.

an electron at a distance δx from the line along which, by assumption, B_x vanishes is given by Eq. (1.17) as $F_a = ev_z(\partial B_x / \partial x) \delta x$. Taking

¹ G. H. Uhlenbeck and S. A. Goudsmit, *Physica*, **5**, 266 (1925); *Nature*, **117**, 264 (1926).

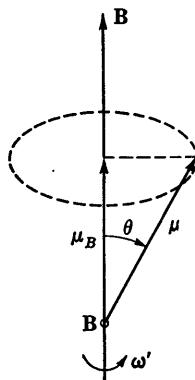


FIG. 2.20. Precession of the moment μ with a constant projection μ_B .

arising from the interaction between the electron's magnetic moment and an inhomogeneity in the magnetic induction. This can be seen from Fig. 2.21 in the following way. Assume an electron to move with a velocity v along the z axis through a magnetic induction B , principally in the y direction, which does not vary with z but has a variation with y and, in consequence of $\nabla \cdot \mathbf{B} = 0$, a variation in x . The necessary relation is $\partial B_y / \partial y = -\partial B_x / \partial x$. The force exerted by B_y on the moment is $F_\mu = \mu_{sB} \partial B_y / \partial y$. The amperian force in the y direction which is experienced by

$\mu_{sB} = e\hbar/2m$, $mv_z = \hbar k_z$, and $\delta x \delta k_x = 1$, the ratio of the forces in the y direction due to the magnetic moment and due to inevitable translational motion is

$$\frac{F_\mu}{F_a} = \frac{\mu_{sB}}{ev_z \delta x} = \frac{\hbar}{2mv_z \delta x} = \frac{1}{2} \frac{1}{k_z \delta x} = \frac{1}{2} \frac{\delta k_x}{k_z}$$

Therefore, as F_μ must be large compared to F_a if the former is to be measured accurately, the uncertainty in k_z , or the uncertainty in the velocity in the x direction, must be large compared to k_z , or the velocity in the z direction. In the postulated beam along z the converse of this is true. Hence the experiment is infeasible.

Recently Crane¹ and his colleagues have succeeded in evaluating ω' by measuring the angle through which the axis of an electron's intrinsic angular momentum precesses while the electron is traversing a measured distance with a known velocity parallel to a magnetic induction of known magnitude. The electron velocity v is obtained from the accelerating potential, and, if d is the path length, B the magnetic induction, and ϕ the angle of precession, ω' is given by $\phi v/d$ and γ by $-v\phi/Bd$. The difficulty is to devise a method of measuring ϕ , since in an ordinary homogeneous beam of electrons the intrinsic-angular-momentum axis may be thought of as uniformly disposed in azimuth about the beam axis. However, if such a beam is scattered (nuclear scattering) at right angles by a solid block of matter, an asymmetry in this azimuthal distribution (polarization) is brought about which can be detected by the dependence of scattered-beam intensity after a second right-angle scattering process. This asymmetry acts as the fiducial azimuth vector, and its rotation about the beam axis measures the precession ϕ in a magnetic induction maintained between the two blocks of scattering material. The precession in the preliminary experiments was not great, but γ was determined to be equal to $(1.99 \pm 0.02)e/2m_0$. Extensive observations and precise measurements have also been made on electrons in atomic systems, as will be discussed in Chap. 4. Such measurements are likewise in accord with the assignment to an electron of an intrinsic angular momentum, or *spin*, and a magnetic moment equal to about e/m times this spin.

There is a second property of these moments which is anomalous classically but which is of basic importance in atomic mechanics. The magnetic moment of an electron cannot be observed at an arbitrary angle with the ambient magnetic induction. It can be observed at but two possible inclinations or can have but two components in the direction of the ambient magnetic induction. The component of the angular momentum can be $+\hbar/2$ or $-\hbar/2$ only, and in consequence the magnetic-

¹ H. R. Crane, R. W. Pidd, and W. H. Louisell, *Phys. Rev.*, **91**, 475 (1953).

moment components can take only the values $\pm e\hbar/2m_0$. The forces and torques that result from the interaction of the magnetic moment with the magnetic induction depend only upon the component of the moment along the magnetic induction. The limitation of permissible orientations of the electron moment to those having these two values of the components along \mathbf{B} instead of permitting the infinite number of possible directions in space is, like the nonclassical value of γ_s , a basic quantum-mechanical postulate, the justification for which is to be found in the adequacy of the resulting description of atomic phenomena. The first direct evidence for this effect was found by Stern and Gerlach¹ when investigating the lateral deflection experienced by a narrow beam of neutral silver atoms. These atoms were evaporated from a furnace and collimated in such a way that they traversed a long path through a region of magnetic induction which had a strong gradient normal to the beam. The magnetic moment of a silver atom is essentially that of an electron (see Sec. 4.5), and in accordance with the above discussion it should experience a force in the direction of the gradient of the magnetic induction which is proportional to the component of the moment along this gradient. Stern and Gerlach found that the displacement of the beam of atoms was not a uniform spreading but that the beam was split into two components, one of which moved toward the region of greater B and the other toward the region of lesser B . This indicated strongly that the orientations of the moments relative to \mathbf{B} were limited to two values, and subsequent, more precise experiments which will be described in later sections have completely confirmed this view.

The relativistic theory of the electron due to Dirac² (Sec. 4.2) is consistent with the assignment to an electron of an intrinsic-angular-momentum component of $\hbar/2$ and a magnetic-moment component of exactly $e\hbar/2m_0$. Thus the value of the ratio, γ_s , would be e/m_0 , exactly twice the classical value of Eq. (2.20). Very precise measurements of this ratio, which have recently been made by the study of the fine structure of energy levels in various atomic systems,³ show that γ_s is not exactly twice γ_l but is somewhat greater than $2\gamma_l$. The proper quantum-mechanical explanation of this fact has been given by Schwinger,⁴ who has shown that to a first approximation the component of the magnetic

¹ O. Stern and W. Gerlach, *Ann. Physik*, **74**, 673 (1924); *Z. Physik*, **41**, 563 (1927).

² P. A. M. Dirac, "The Principles of Quantum Mechanics," 2d ed., Oxford University Press, London, 1935.

³ J. E. Nafe and E. B. Nelson, *Phys. Rev.*, **73**, 718 (1948); P. Kusch and H. M. Foley, *Phys. Rev.*, **74**, 250 (1948); S. Koenig, A. G. Prodell, and P. Kusch, *Phys. Rev.*, **83**, 687 (1951).

⁴ J. Schwinger, *Phys. Rev.*, **73**, 416 (1948); **74**, 2439 (1948); **75**, 651 (1949); **76**, 790 (1950).

moment associated with electron spin along the direction of the local magnetic induction is given by¹

$$\mu_{eB} = \pm \left(1 + \frac{\alpha}{2\pi}\right) \frac{e\hbar}{2m_0} \quad (2.23)$$

where α , which is the fine-structure constant of Sec. 2.4, is given by $e^2/4\pi\epsilon_0ch$. This is in excellent agreement with the experimental values. The derivation of the above relationship is beyond the scope of this treatment, but, on the assumption that the relativistic theory of Dirac, as outlined in Sec. 4.2, is correct in accounting for the anomalous value for the ratio γ_s , it can be seen qualitatively from simple classical considerations that when measured in units of $e\hbar/2m_0$ the spin moment should appear to be greater than unity by a quantity of the order of magnitude of α . This may be thought of as due to the fact that the value of the effective inertial mass of an electron, which is the m_0 of the ratio $e\hbar/2m_0$, is greater than its effective mass for spinning motion. As a consequence of the uncertainty jittering or pulsation, the electron may be thought of classically as exchanging radiant energy with its surroundings, whereas in its intrinsic spinning motion, since there is no linear acceleration, no radiation is to be associated with the motion. Since the mass occurs in the denominator of the equation for the magnetic moment, the value of the spin moment in terms of the inertial mass would be expected to be larger than if the mass had been measured in terms of the mass involved in the spin. This situation may be thought of as somewhat analogous to the apparent increase in the mass of a vibrating diaphragm or string when surrounded by air which can radiate away the acoustic energy of the vibrating system. The magnitude of the effect to be expected in the case of the electron can be seen qualitatively from the argument leading to Eq. (1.25). If the radiation damping force F_z is set equal to an effective radiation mass correction δm times the acceleration \ddot{z} ,

$$\delta m \ddot{z} = \frac{e^2}{6\pi\epsilon_0 c^3} \ddot{z}$$

If z is simply periodic, with a frequency corresponding to an uncertainty time of the Compton wavelength divided by c , that is, $mc^2 = h\nu$, then the magnitude of \ddot{z} is equal to $2\pi\nu$ times the magnitude of \dot{z} , or

$$\delta m = \frac{e^2}{6\pi\epsilon_0 c^3} \frac{mc^2}{\hbar}$$

¹ R. Karplus and N. Kroll, *Phys. Rev.*, **77**, 536 (1950), have calculated the parenthesis to the next higher order in α and found that it is $1 + \alpha/2\pi + 2.973\alpha^2/\pi^2$.

and

$$\frac{\delta m}{m} = \frac{2}{3} \frac{e^2}{4\pi\epsilon_0 c\hbar} = \frac{2}{3}\alpha$$

This simple argument yields the correct order of magnitude for the correction term of the mass due to associated radiation, but, since the classical theory of radiation is not adequate for such a calculation, the result is about four times too great.

Atomic Nuclei and Their Constituents. Early observations of the hyperfine structure of atomic spectra led to the conclusion that atomic nuclei also possess integral multiples of the angular momentum found for the electron spin. The integer associated with a nucleus can in many cases be determined from molecular spectra and hyperfine structure in atomic spectra (Sec. 4.4). From a knowledge of the nuclear angular momentum and the assumption that for any available values of \mathbf{B} the mechanical and magnetic moments remain collinear and in a constant ratio to one another, the magnetic moment of an atomic nucleus can be found from a determination of γ by measuring ω' and \mathbf{B} [Eq. (2.22)]. The values found for the moments are as specific nuclear characteristics as are masses and angular momenta. They are smaller by a factor of the order of 10^3 than the electron-spin magnetic moment. The values of γ for atomic nuclei or their constituents which are known as *nucleons* are not even approximated by any multiple of one-half their charge-to-mass ratio. This indicates clearly a fundamental and as yet little-understood difference between electrons and nucleons.

Two general techniques have been employed for the precision determination of nuclear magnetic moments. The first of these, used by Rabi and his colleagues, employs a molecular beam from which the atoms or molecules are removed by inducing a reorientation of their nuclear moments in a uniform \mathbf{B} under the influence of an electromagnetic field of a frequency resonant with that of the Larmor precessional frequency of the nucleus in the presence of \mathbf{B} . This type of experiment will be considered further in Secs. 3.3 and 4.4. The second technique employs matter in bulk which is placed in a region of uniform \mathbf{B} , and the resonant condition between the Larmor frequency of nuclear precession and the frequency of the applied electromagnetic field or magnetic induction is observed. The conditions that must be fulfilled to obtain a sharp, observable resonance condition are that the nuclear systems under observation should be (1) free enough from orienting forces exerted on them by surrounding electrons and neighboring atoms to be able to precess in response to the torque applied to them by the varying electromagnetic field, and (2) not completely free of such forces, as it is necessary that the systems should be able to transfer the energy they gain from the exciting

circuit to the crystal lattice. Otherwise they would be unable to accept energy indefinitely from the circuit, and hence no reaction upon the circuit at resonance could be observed. These conditions are sufficiently well fulfilled by nuclei to enable the resonant condition to be detected, and a further study of the shapes and structures of the resonant-response curves furnishes valuable information on the nature of the torques acting between nuclei and their atomic surroundings. By this means, a number of quantities of atomic interest can be studied.¹

Two general methods have been employed for the detection of the resonant condition in matter. One of these, used by Purcell and Pound and their groups,² detects the resonant condition by a bridge measurement of the complex impedance as a function of frequency of an electric circuit element near the material under study. The characteristics of the atomic nuclei in the material near or enclosed by the circuit element

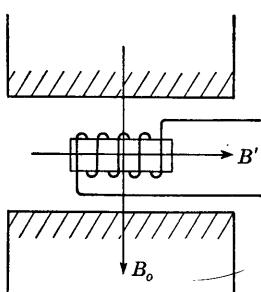


FIG. 2.22. Magnetic-resonance absorption.

can be related to the observed impedance. The maximum power absorption may be thought of as occurring at the closest possible approach to the ideal condition of synchronous resonant precession, in which all the nuclear moments precess in phase with one another about the principal magnetic induction B_0 under the influence of the alternating magnetic vector B' in the coil of Fig. 2.22. As the precession is not completely free but is under the influence of the forces exerted by the surrounding lattice atoms in the case of a crystalline solid, the energy of the pre-

cessing system is constantly being lost to the lattice. The way in which the oscillating magnetic induction B' supplies power to the precessing nuclear moments may be thought of in terms of a complex susceptibility of the aggregate of nuclear moments. The inductance of the coil may be written $L = L_0(1 + f\chi)$, where L_0 would be the inductance in the absence of the enclosed material, f is a geometrical factor relating to the fraction of the space near the coil occupied by the material, and χ , the susceptibility, is the induced magnetic moment per unit volume per unit magnetic induction B' . If there were no time lag in the establishment of the magnetization, or magnetic moment along B' , χ would be real, but, as the moment results both from B' and from the

¹ B. V. Rollin, *Repts. Progr. in Phys.*, **12**, 22 (1949); N. Bloembergen, E. M. Purcell, and R. V. Pound, *Phys. Rev.*, **73**, 679 (1948); G. E. Pake, *Am. J. Phys.*, **18**, 438, 473 (1950); J. H. Van Vleck, *Am. J. Phys.*, **18**, 495 (1950); M. A. Garstens, *Phys. Rev.* **81**, 645 (1951).

² E. M. Purcell, H. C. Torrey, and R. V. Pound, *Phys. Rev.*, **69**, 37 (1946); E. M. Purcell, N. Bloembergen, and R. V. Pound, *Phys. Rev.*, **70**, 988 (1946); R. V. Pound and W. D. Knight, *Rev. Sci. Instr.*, **21**, 219 (1950) (circuitry).

opposing effects of lattice forces, there is a phase lag. In the usual complex circuit notation this implies that $\chi = \chi' + i\chi''$, where χ'' is the out-of-phase component. The impedance of the coil terminals is then

$$z = R + i\omega L = (R - \omega L_0 f\chi'') + i\omega L_0(1 + f\chi')$$

Thus χ'' contributes to the real part of the impedance. This is a maximum when ω is the same as the precessional frequency in the magnetic induction B_0 , and the sharpness of resonance and maximum value of energy absorption obtained depend upon the nature and strength of the lattice forces.

The other general method of determining the resonant condition, used by Bloch and Hansen and their group,¹ is closely the same in principle and may be understood in terms of the preceding description. A second coil is so disposed around the sample that its axis is at right angles to both B_0 and B' . It is evident from the previous discussion that, as the frequency of B' approaches the resonant frequency of precession of the nuclear moments about B_0 , these will tend to precess in phase with one another about B_0 and give rise to an oscillating component of magnetic moment through the detecting coil at right angles to B_0 and B' (see Fig. 2.23). The imaginary component of the susceptibility, spatially normal to B' , is just the magnetic moment per unit volume per unit B' in the direction of the axis of the detecting coil, and a maximum signal appears at the terminals of this coil for the resonant condition between the frequency of B' and that of nuclear precession. For further details the references should be consulted.

From a knowledge of ω' and the measured value of B , which may have to be corrected for magnetic effects of the surrounding electron structure or neighboring atoms, the value of the ratio of the magnetic moment to the angular momentum can be obtained from Eq. (2.22). If I is the angular momentum of the nucleus in units of \hbar , $\mu = \gamma I \hbar$, and, since $\gamma = -\omega'/B$,

$$-\mu = \frac{\omega' I \hbar}{B} = \left(\frac{2\omega' m_0 I}{Be} \right) \frac{e\hbar}{2m_0} \quad (2.24)$$

$$-\mu = \left(\frac{2\omega' M_p I}{Be} \right) \frac{e\hbar}{2M_p} \quad (2.24')$$

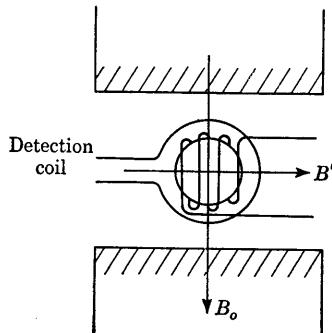


FIG. 2.23. Nuclear-resonance induction.

¹ F. Bloch, W. W. Hansen, and M. Packard, *Phys. Rev.*, **69**, 127 (1946); **70**, 474 (1946); F. Bloch, *Phys. Rev.*, **70**, 460 (1946).

The quantities in parentheses are the values of the moments in terms of the *Bohr magneton* ($e\hbar/2m_0$) and the *nuclear magneton* ($e\hbar/2M_p$), where m_0 and M_p are the electron rest mass and proton mass, respectively. It is easier to make precise relative measurements in which B need not be known accurately, and, evidently, relative values of μ may be obtained by placing the samples in the same magnetic induction and measuring frequencies only, if the values of I are known. Thus

$$\frac{\mu_1}{\mu_2} = \frac{\omega'_1 I_1}{\omega'_2 I_2}$$

The values of I can in most instances be found from spectroscopy (Secs. 4.4 and 5.5). Experiments of this kind have been used to determine the relative values of μ for the proton and neutron^{1,2} ($I = \frac{1}{2}$ in both cases) and for the proton and deuteron^{1,3} ($I_D = 1$). In a similar way such ratios have been observed^{4,5} for H^3 and He^3 . Many other comparative measurements of this type have been made, particularly by R. V. Pound⁶ and F. Bitter,⁷ and are the basis for the values of the nuclear magnetic moments to be found in the tables.

$$\begin{aligned}\frac{\mu_n}{\mu_p} &= -0.685001 \pm 0.00003 \dagger \\ \frac{\mu_D}{\mu_p} &= 0.3070122 \pm 0.0000001 \ddagger \\ \frac{\mu_T}{\mu_p} &= 1.06666 \pm 0.00010 \\ \frac{\mu_{He^3}}{\mu_p} &= 0.763 \pm 0.007\end{aligned}$$

Certain other important comparative experiments have been performed by Gardner and Purcell⁸ and by Taub and Kusch.⁹ Taub and Kusch compared the proton precession frequency ω'_p with the frequency

¹ W. R. Arnold and A. Roberts, *Phys. Rev.*, **71**, 878 (1947).

² F. Bloch, D. Nicodemus, and A. H. Staub, *Phys. Rev.*, **74**, 1025 (1948).

³ F. Bloch, E. C. Leventhal, and M. E. Packard, *Phys. Rev.*, **72**, 1125 (1947).

⁴ H. L. Anderson and A. Novick, *Phys. Rev.*, **71**, 372 (1947); F. Bloch, A. C. Graves, M. E. Packard, and R. W. Spence, *Phys. Rev.*, **71**, 373 (1947).

⁵ H. L. Anderson and A. Novick, *Phys. Rev.*, **73**, 919 (1948).

⁶ R. V. Pound, *Phys. Rev.*, **72**, 527, 1273 (1947); **73**, 523, 1112, 1247 (1948).

⁷ F. Bitter, *Phys. Rev.*, **75**, 1326 (1949).

[†] Sign (proton +, neutron -) from H. H. Staub and E. H. Rogers, *Phys. Rev.*, **76**, 980 (1949).

[‡] B. Smabler, E. Yasaitis, and H. L. Anderson, *Phys. Rev.*, **81**, 896 (1951).

⁸ J. H. Gardner and E. M. Purcell, *Phys. Rev.*, **76**, 1262 (1949); J. H. Gardner, *Phys. Rev.*, **83**, 996 (1951).

⁹ A. H. Taub and P. Kusch, *Phys. Rev.*, **75**, 148 (1949).

of certain hyperfine-structure transitions in Cs^{133} and In^{115} using the molecular-beam technique which effectively measures ω'_p in terms of the electron-spin precession frequency. Using Eq. (2.23) to interpret μ_s in terms of $e\hbar/2m_0$, the proton moment is found to be

$$\begin{aligned}\mu_p &= (1.51929 \pm 0.00008) \times 10^{-3} \mu_s \\ &= (1.52106 \pm 0.00008) \times 10^{-3} \frac{e\hbar}{2m_0}\end{aligned}$$

Gardner and Purcell have measured the ratio of ω'_p to the electron cyclotron frequency ω_e directly, which by Eqs. (2.7) and (2.24) yields $\mu_p = \omega'_p e\hbar / \omega_e 2m_0$, since $I_p = \frac{1}{2}$ for a proton. Figure 2.24 illustrates schematically the way in which the experiment was performed. The oscillating electromagnetic field was produced by a standing microwave in an evacuated wave guide. A magnetic induction \mathbf{B}_0 was produced normal to the smaller surfaces of the guide, and the proton resonant condition was determined in the usual way by observing the frequency for maximum power absorption from the electromagnetic wave by the protons in a sample of mineral oil in the guide. In the neighborhood of the sample, effectively the same \mathbf{B}_0 , two slits in the walls of the guide permitted a beam of electrons to pass from a heated filament on one side, across the evacuated guide in the direction of \mathbf{B}_0 , and to a collector on the far side of the guide. When the variable frequency of \mathbf{E}' nears ω_e the resonant condition tends to increase the amplitude of the helical motion of the electrons about \mathbf{B}_0 and this affects the number that can pass eventually through the second slit and reach the collector beyond it. Thus the variation of the collector current as the frequency of \mathbf{E}' is varied permits the determination of ω_e . By a precise method of harmonic comparison it was found that $\omega_e/\omega'_p = 657.475 \pm 0.008$. Making a small correction for diamagnetic shielding of the proton by the surrounding electron in the hydrogen atom, it was found that

$$\mu_p = (1.52100 \pm 0.00002) \times 10^{-3} \frac{e\hbar}{2m_0}$$

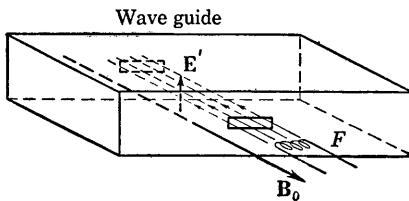


FIG. 2.24. Schematic of Gardner and Purcell apparatus for measuring the electron cyclotron frequency.

This is in excellent agreement with the results of Taub and Kusch.

Hipple¹ and his associates have made a precise measurement of the

¹ H. A. Thomas, R. L. Driscoll, and J. A. Hipple, *Phys. Rev.*, **75**, 902 (1949); H.

ratio of the precessional frequency to the cyclotron frequency for the proton (see also Bloch and Jeffries¹). By Eqs. (2.7) and (2.22) this yields

$$\frac{\omega'_p}{\omega_p} = \frac{M_p \gamma}{e} = \mu_p \left(\frac{e\hbar}{2M_p} \right)^{-1}$$

where μ_p is the proton magnetic moment in units of nuclear magnetons. The apparatus used by Hipple for measuring the cyclotron frequency ω_p was described in Sec. 2.3, and the precessional frequency was measured by the standard absorption technique for a proton-containing material in the same region of magnetic induction. The value obtained for μ_p was

$$\mu_p = 2.792685 \pm 0.00003 \frac{e\hbar}{2M_p}$$

From the value of e/M_p obtained from ω_p and B , the value of γ_{sp} for the spinning motion of the proton is obtained:

$$\gamma_{sp} = (2.67530 \pm 0.00006) \times 10^8 \text{ m}^2 \text{ weber}^{-1} \text{ sec}^{-1}$$

This value of $\gamma_{sp} = \mu_p/(\hbar/2)$ can be combined with Gardner and Purcell's value of μ_p in units of $e\hbar/2m_0$ to yield

$$\begin{aligned} \frac{e}{m_0} &= \frac{2.67530 \pm 0.00006}{1.52100 \pm 0.000002} \times 10^{11} \\ &= (1.75891 \pm 0.00005) \times 10^{11} \text{ coulombs kg}^{-1} \end{aligned}$$

Also, to the precision with which \hbar is known, the absolute value of the moment is

$$\mu_p = (1.41012 \pm 0.00005) \times 10^{-26} \text{ joule-m}^2 \text{ weber}^{-1}$$

A least-squares adjusted value² is slightly higher:

$$\mu_p = (1.41045 \pm 0.00009) \times 10^{-26} \text{ joule-m}^2 \text{ weber}^{-1}$$

It is interesting, finally, to note that the ratio of the moment in nuclear magnetons to the value in Bohr magnetons is the ratio of the mass of the proton to that of the electron. Forming this quantity,

$$\frac{M_p}{m_0} = \frac{2.79268 + 0.00006}{1.52100 \pm 0.00002} \times 10^3 = 1836.08 \pm 0.05$$

Sommer, H. A. Thomas, and J. A. Hipple, *Phys. Rev.*, **80**, 487 (1950); **82**, 697 (1951); H. A. Thomas, *Phys. Rev.*, **80**, 901 (1950).

¹ F. Bloch and C. D. Jeffries, *Phys. Rev.*, **80**, 305 (1950); C. D. Jeffries, *Phys. Rev.*, **81**, 1040 (1951).

² J. W. M. DuMond and E. R. Cohen, *Revs. Mod. Phys.*, **25**, 691 (1953).

This is in excellent agreement with the value determined by other methods and provides an accurate value for this important atomic quantity.

Photons. In Sec. 2.4 it was seen that the phenomena of radiation could be described in terms of the concept of particles having zero rest mass and the constant velocity c in free space, known as photons. They are analogous to material particles only in a limited sense for there is no nonrelativistic approximation applicable to photons. Like material particles, their trajectories may be thought of as guided by a wave equation which in the case of photons is derivable from Maxwell's equations. In contrast to material particles, which in the limit of very great masses and high associated frequencies would be measured with a balance, photons of very low frequency are measured in terms of their associated electric and magnetic vectors by instruments such as voltmeters and ammeters. The concepts of energy, momentum, and angular momentum are applicable to photons as well as to particles with finite rest masses. As was seen in Sec. 1.8 a dipole oscillator in the form of a charge rotating in a circle about an equal charge of unlike sign radiates angular momentum at $1/\omega$ times the rate it radiates energy. Thus a photon of energy $\hbar\omega$ radiated by such a dipole would carry with it the angular momentum \hbar . Unlike the angular momentum associated with a particle of finite rest mass, this cannot be analyzed in terms of intrinsic spin and extrinsic angular momentum with respect to the emitting atom. It will be seen in Sec. 3.6 that in a dipole-radiation process an atom's angular momentum changes by \hbar , corresponding closely to the analogy of a rotating charge. A photon with the angular momentum \hbar is, however, to be distinguished from a material particle with a spin \hbar , in that two senses of plane or circular polarization are possible for the photon whereas for a material particle spatial quantization would lead to three possible orientations of the spin vector. Again, as the photon has no charge associated with it, no magnetic moment would be anticipated nor is one observed.

2.7. Higher-order Moments. The electric monopole moment represented by the nuclear electric charge and the magnetic dipole moment discussed above are merely the first approximations to the electric and magnetic moments of a general charge distribution in arbitrary motion. Interest in atomic and nuclear physics centers very largely on charge distributions in simple rotational motion, and the concept of higher moments can be more easily exhibited here than in the general case. Observations of the behavior of atomic and nuclear systems are most directly correlated with their energy. The energy of a charge e in a region where the potential due to external charges is φ , is $e\varphi$. The energy of a charge rotating with an angular velocity ω about an area a in a region of magnetic induction \mathbf{B} was earlier seen to be

$$U_m = -\mathbf{u} \cdot \mathbf{B} = -(ea/2\pi)\omega \cdot \mathbf{B}$$

This can be written more conveniently by applying Stokes's theorem, $\int_l \mathbf{A} \cdot d\mathbf{l} = \int_a \nabla \times \mathbf{A} \cdot d\mathbf{a}$, to the circulating charge and using $\mathbf{B} = \nabla \times \mathbf{A}$ to obtain $2\pi r' A_\psi = aB$, from which $U_m = -er' A_\psi \omega$. Here r' is the radius of the area a , and A_ψ is the component of \mathbf{A} about the polar axis of rotation. Thus the total electric and magnetic energy of the rotating charge distribution is given by

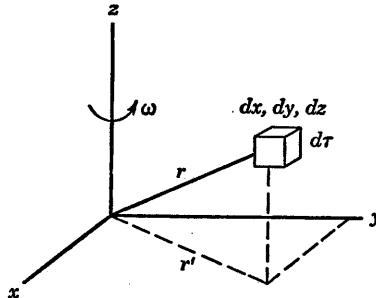


FIG. 2.25. Charge-density integration.

$$U = \int \rho \varphi d\tau - \omega \int \rho r' A_\psi d\tau$$

where $\rho d\tau$ is the charge in the infinitesimal volume element $d\tau$.

In general, ρ , φ , and A_ψ are all functions of the coordinates. An arbitrary distribution of charge ρ can be built up of two distributions ρ_e and ρ_o , where these are even functions in the sense that

$$\rho_e(x, y, z) = \rho_e(-x, -y, -z)$$

and odd in the sense that $\rho_o(x, y, z) = -\rho_o(-x, -y, -z)$, respectively. It will be assumed that φ and A_ψ are such that they can be expanded in a Taylor series about the origin, which is also assumed to be the center of mass about which the charges rotate (see Fig. 2.25). Then

$$\varphi = \varphi_0 + \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + z \frac{\partial}{\partial z} \right)_0 \varphi + \frac{1}{2} \left(\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} \right)_0 \varphi + \dots$$

and similarly for A_ψ . The energy thus becomes

$$\begin{aligned} U = & \underbrace{\varphi_0 \int (\rho_e + \rho_o) d\tau}_{1} - \underbrace{\omega A_{\psi 0} \int (\rho_e + \rho_o) \sqrt{x^2 + y^2} d\tau}_{2} \\ & + \underbrace{\left(\frac{\partial \varphi}{\partial x} \right)_0 \int (\rho_e + \rho_o) x d\tau + \left(\frac{\partial \varphi}{\partial y} \right)_0 \int (\rho_e + \rho_o) y d\tau + \left(\frac{\partial \varphi}{\partial z} \right)_0 \int (\rho_e + \rho_o) z d\tau}_{3} \\ & - \underbrace{\left[\left(\frac{\partial A_\psi}{\partial x} \right)_0 \omega \int (\rho_e + \rho_o) x \sqrt{x^2 + y^2} d\tau + \dots + \dots \right]}_{4} \\ & + \underbrace{\frac{1}{2} \left(\frac{\partial^2 \varphi}{\partial x^2} \right)_0 \int (\rho_e + \rho_o) x^2 d\tau + \dots}_{5} \end{aligned}$$

These terms are known as the energies related to the electric monopole (1), dipole (3), quadrupole (5), etc., and the magnetic dipole (2), quad-

rupole (4), etc. The integrals are known as the moments associated with the several multipoles. If the change in the potentials over the volume occupied by the charge distribution is small, successive terms decrease in magnitude. Also half the terms can be dropped because of symmetry. From the definitions of ρ_e and ρ_o , the contributions to the integrals from symmetrically placed volume elements vanish in pairs for one or the other of these depending on whether an odd or even function of the coordinates appears with it in the integrand. Thus

$$\begin{aligned}
 U = & \underbrace{\varphi_0 \int \rho_e d\tau - A_{\psi_0} \omega \int \rho_e \sqrt{x^2 + y^2} d\tau}_{1} \\
 & + \underbrace{\left(\frac{\partial \varphi}{\partial x} \right)_0 \int \rho_o x d\tau + \dots + \dots}_{3} \\
 & - \underbrace{\left[\left(\frac{\partial A_\psi}{\partial x} \right)_0 \omega \int \rho_o x \sqrt{x^2 + y^2} d\tau + \dots + \dots \right]}_{4} \\
 & + \underbrace{\frac{1}{2} \left(\frac{\partial^2 \varphi}{\partial x^2} \right)_0 \int \rho_e x^2 d\tau + \left(\frac{\partial^2 \varphi}{\partial x \partial y} \right)_0 \int \rho_e xy d\tau + \dots}_{5}
 \end{aligned}$$

Thus the electric monopole moment, magnetic dipole moment, electric quadrupole moment, etc., arise from even distributions of charge only. The other terms arise from odd distributions. Both types occur and are of interest in rotating molecules. But the elementary particles dealt with in this section are such that their charge distributions are found to be even. Thus nuclei of atoms have electric monopole moments (a net charge), magnetic dipole moments, and electric quadrupole moments. Successively higher nonvanishing moments in the two series, electric and magnetic, become smaller by a factor of the order of the square of nuclear dimensions. It should also be noticed that the contributions to the energy made by the series of moments result from their products with successively higher orders of differentials of the potentials. Thus the more sharply these potentials vary, the greater is the energy contribution from higher moments.

The electric quadrupole moment is the only one besides the electric monopole moment and magnetic dipole moment that has so far been observed with any assurance for nuclei. The magnetic dipole and electric quadrupole moments are the nuclear properties primarily responsible for the hyperfine structure of spectra (Sec. 4.4). The quadrupole moment is a quantity as characteristic of a particular nucleus as are

TABLE 2.1. SUMMARY OF PROPERTIES OF ELECTRONS, NUCLEONS, AND PHOTONS

Property	Electron	Nucleon	Photon
Mechanical:			
Rest mass	m_0	M_0	0
Momentum	$m\mathbf{v} = \hbar\mathbf{k}$	$M\mathbf{v} = \hbar\mathbf{k}$	$\frac{\hbar\omega}{c} = \hbar k$
Energy	$mc^2 = \hbar\omega$	$Mc^2 = \hbar\omega$	$\hbar\omega$
Intrinsic angular-momentum component	$\hbar/2$	$\hbar/2$	\hbar (dipole)
Electric and Magnetic:			
Monopole moment	$-e$ (electron) $+e$ (positron)	$+e$ (proton) 0 (neutron) Ze (nucleus)	0
Intrinsic magnetic-dipole-moment component	$\sim \left(1 + \frac{\alpha}{2\pi}\right) \mu_0$	$g\mu_p$	0
Electric quadrupole-moment magnitude	0	Q	0

$$\omega = ck, \quad \mu_0 = e\hbar/2m_0, \quad \alpha = e^2/4\pi\epsilon_0\hbar c$$

Z and g characteristic dimensionless constants.

Q characteristic constant, dimension of length squared.

charge, mass, spin, and magnetic dipole moment. It may be that the existence of an electric quadrupole moment is an indication of complexity

of structure, for no such moments have been observed for protons or electrons but only for a nuclear spin equal to or greater than 1.

Finally, as it has been assumed that the charge distribution is in rotation, there is on the average an axial symmetry about z, and the expression for the quadrupole moment can be somewhat further simplified: x and y may be written $r' \cos \omega t$ and $r' \sin \omega t$, respectively, and the cross-product quadrupole

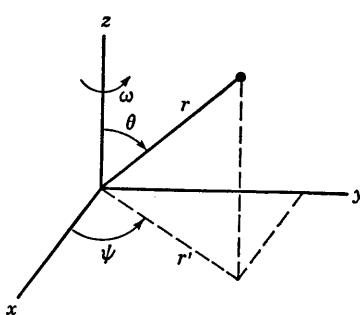


FIG. 2.26. Quadrupole integration.

terms vanish because ρ_e has then the additional properties $\bar{\rho}_e(x,y,z) = \bar{\rho}_e(-x,y,z)$ and $\bar{\rho}_e(x,y,z) = \bar{\rho}_e(x,-y,z)$. The bar indicates the average value.

$$U_Q = \frac{1}{2} \left[\left(\frac{\partial^2 \varphi}{\partial x^2} \right)_0 \int (\cos^2 \omega t) r'^2 \rho \, d\tau + \left(\frac{\partial^2 \varphi}{\partial y^2} \right)_0 \int (\sin^2 \omega t) r'^2 \rho \, d\tau \right. \\ \left. + \left(\frac{\partial^2 \varphi}{\partial z^2} \right)_0 \int z^2 \rho \, d\tau \right]$$

The time-average value of sine squared or cosine squared is $\frac{1}{2}$, and, since $\nabla^2 \varphi = 0$,

$$\overline{U}_Q = \frac{1}{2} \left(\frac{\partial^2 \varphi}{\partial z^2} \right)_0 \int (z^2 - \frac{1}{2} r'^2) \bar{\rho} d\tau$$

and, since $r^2 = r'^2 + z^2$ and $z = r \cos \theta$, as in Fig. 2.26,

$$\overline{U}_Q = \frac{1}{4} \left(\frac{\partial^2 \varphi}{\partial z^2} \right)_0 \int r^2 (3 \cos^2 \theta - 1) \bar{\rho} d\tau$$

The quadrupole moment Q is generally defined as this integral divided by the total charge, which is a measure of the ellipticity of the charge distribution. If $\bar{\rho}$ is small for large θ , the integral is positive and the distribution is prolate or cigar-shaped. If $\bar{\rho}$ is large for large θ , the integral is negative and the distribution is oblate or discus-shaped. A few representative quadrupole moments are given in Table 2.2. Mag-

TABLE 2.2. CHARACTERISTIC PARAMETERS FOR ELEMENTARY PARTICLES AND LIGHT NUCLEI¹

Particle	Atomic mass, in physical atomic mass units. Error in parts per 10^6 in parentheses	Nuclear charge, in units of e	Nuclear angular momentum, in units of \hbar	Nuclear magnetic moment, in units of $e\hbar/2M_p$ (except for e)	Nuclear quadrupole moment, in units of 10^{-28} m^2
Electron (e)	5.48749×10^{-4} (18)	-1	$\frac{1}{2}$	$1.00119(50) \mu_0$	0
Neutron (n^0)	1.008982 (3)	0	$\frac{1}{2}$	$-1.91280(9) \mu_p$	0
H ¹	1.008142 (2)	1	$\frac{1}{2}$	2.79255	0
H ²	2.014735 (2)	1	1	0.857348 (3)	+0.002739
He ³	3.016977 (4)	2	$\frac{1}{2}$	-2.127414 (0.3)	
He ⁴	4.003873 (4)	2	0	0	0
Li ⁶	6.017021 (4)	3	1	0.82189 (40)	≤ 0.0005
Li ⁷	7.018223 (4)	3	$\frac{3}{2}$	3.25586 (11)	+0.02
Be ⁹	9.015043 (3)	4	$\frac{3}{2}$	-1.1774 (80)	
B ¹⁰	10.016114 (3)	5	3	1.8004 (70)	+0.06
B ¹¹	11.012789 (2)	5	$\frac{3}{2}$	2.68858 (10)	+0.03
C ¹²	12.003804 (1.5)	6	0	0	
C ¹³	13.007473 (1)	6	$\frac{1}{2}$	0.70225 (150)	
N ¹⁴	14.007515 (1)	7	1	0.40365 (10)	+0.02
N ¹⁵	15.004863 (1)	7	$\frac{1}{2}$	(\pm) 0.28299 (10)	
O ¹⁶	16.	8	0	0	
O ¹⁷	17.004533 (0.5)	8	$\frac{5}{2}$	-1.8928 (101)	-0.005
O ¹⁸	18.004874 (1)	8	0		

¹ Data from: *Natl. Bur. Standards Circ.* 499, Nuclear Data, 1950 and Supplement; C. W. Li, W. Whaling, W. A. Fowler, and C. C. Lauritsen, *Phys. Rev.*, **83**, 512 (1951); **88**, 1083 (1952); H. Ewald, *Z. Naturforsch.*, **69**, 293 (1951); N. F. Ramsey, pt. III, vol. I, of E. Segre, "Experimental Nuclear Physics," John Wiley & Sons, Inc., New York, 1953; K. T. Bainbridge, pt. V, vol. I, of E. Segre, "Experimental Nuclear Physics," John Wiley & Sons, Inc., New York, 1953.

netic dipole moments exist only if the nuclear spin is greater than zero, and quadrupole moments only if it is equal to or greater than 1.

PROBLEMS

2.1. Show that Eq. (2.1) can be derived on the assumption that the suspended particles of mass m' occupy the volume independently of the suspending fluid and that they obey the perfect-gas law, $p = n'kT$, where n' is the number of particles per unit volume.

2.2. If the principal crystal axes l_1 , l_2 , and l_3 are all at right angles to one another, show that $d(n)$ of Eq. (2.5) is given by

$$\left[\left(\frac{n_1}{l_1} \right)^2 + \left(\frac{n_2}{l_2} \right)^2 + \left(\frac{n_3}{l_3} \right)^2 \right]^{-\frac{1}{2}}$$

2.3. Show that the sum of the squares of the right-hand sides of Eq. (2.4) is equal to the reciprocal of the square of the unit separation of the family of parallel crystal planes that have intersections sl_1/n_1 , sl_2/n_2 , sl_3/n_3 on the crystal vector directions l_1 , l_2 , l_3 , respectively. This is most easily done by forming the unit vector perpendicular to these planes which is the vector product of any pair of differences between the three vectors l_1/n_1 , l_2/n_2 , l_3/n_3 divided by the magnitude of this vector. Then the plane separation is the scalar product of any one of these vectors and the unit vector perpendicular to the plane.

2.4. Assuming that the calcium and carbon atoms in calcite are spaced as shown in Fig. 2.7, that d_c is 3.036×10^{-11} m, and that the angle between the crystal axes is $101^\circ 55'$, show that the spacing between calcium and carbon atoms is 3.213×10^{-11} m and that the shortest distance between calcium atoms is 4.048×10^{-11} m.

2.5. If r is the angle between the edges of a calcite crystal, show that the angle s between the normals to the faces of the crystal is given by $2 \sin(s/2) \cos(r/2) = 1$ and hence $s = 105^\circ 3'$. This is the crystallographic angle that is actually most convenient to measure experimentally.

2.6. A region of uniform magnetic induction \mathbf{B} which is free of electric fields is bounded by a plane which is parallel to \mathbf{B} . An ion of charge e and mass m enters the region through a small hole in the plane with a velocity \mathbf{v} making an angle θ with the normal to the plane, and the projection of \mathbf{v} on the plane makes an angle ϕ with \mathbf{B} . Show that the ion will be bent around and strike the plane at a point having the coordinates

$$y = 2\rho_0 \cos \theta, \quad x = 2\rho_0 \sin \theta \cos \phi \cot^{-1}(\tan \theta \sin \phi)$$

where the hole is the origin, x is parallel to \mathbf{B} , y is perpendicular to \mathbf{B} , and $\rho_0 = mv/B_e$. If a stream of ions diverges from the hole within the cone of small semiangle $\delta\theta$ about the normal, show that the equation of the points at which the ions strike the plane is

$$x = \pi \cos \phi (2\rho_0^2 - \rho_0 y)^{\frac{1}{2}}$$

If the ion density is uniform in ϕ , what will be the nature of the current density near $x = 0$, $y = 2\rho_0$?

2.7. Consider the motion of an ion in a region of uniform magnetic induction \mathbf{B} and uniform electric field \mathbf{E} , and write \mathbf{v} of Eq. (1.17) as $\mathbf{V} + \mathbf{v}'$, where \mathbf{V} is given by $\mathbf{V} = \mathbf{E} \times \mathbf{B}/|B|^2$. Show that the motion of the ion can be described in terms of the following three motions: (1) a uniform acceleration in the direction of \mathbf{B} equal to $(e\mathbf{B} \cdot \mathbf{E}/m|B|^2)\mathbf{B}$, (2) a uniform translation with the velocity \mathbf{V} , (3) a circular

motion in the plane normal to \mathbf{B} with an angular velocity $\omega = -(e/m)\mathbf{B}$, and a radius $m\mathbf{B} \times \mathbf{v}'/e|\mathbf{B}|^2$.

2.8. Ions are formed in a region of gas at low pressure by a beam of electrons in the form of a cylinder of very small diameter. The axis of the electron beam is parallel to the uniform magnetic induction \mathbf{B} in the region, and a uniform electric field \mathbf{E} exists perpendicular to \mathbf{B} . Show that the path of the ions thus formed is described by the motion of a point a distance r from the center of a wheel of radius R which rolls in the direction perpendicular to \mathbf{E} and \mathbf{B} with the angular velocity eB/m and which is displaced parallel to its axis with the velocity v_{\parallel} , where $R = mE/eB^2$ and $r = mv_{\perp}/eB$. v_{\parallel} and v_{\perp} are the components of the ions' initial velocity parallel and perpendicular to \mathbf{B} , respectively. Sketch the curves traced out for various values of v_{\perp} and show that the line source of ions is focused in a parallel line displaced a distance $(2\pi mE)/eB^2$ in the direction of the wheel's rolling motion.

2.9. Ions of charge e and momentum mv diverge from a point in a region of uniform magnetic induction \mathbf{B} . If the origin of coordinates is at the point source of ions, the x coordinate is in the direction of \mathbf{B} , and r is a radius vector from this axis, show that the trajectory of an ion leaving the source at an angle α with the x axis is

$$r = 2\rho_0 \sin \alpha \sin \left(\frac{x}{2\rho_0 \cos \alpha} \right)$$

where $\rho_0 = mv/Be$. If a very small hole is made on the x axis in a barrier at a distance d from the source and if there is also a barrier plane normal to the x axis at $x = d/2$ in which there is a narrow annular opening of radius S concentric with the x axis, show that the momentum of ions able to pass through the ring and enter the small hole is

$$p = \frac{eB}{2n\pi} (d^2 + \pi^2 n^2 S^2)^{\frac{1}{2}}$$

where n is an odd integer.

2.10. An apparatus for measuring the ratio of the charge of an electron to its mass consists of a long cylindrical tube at one end of which is a filament for liberating electrons thermionically and at the other end is an electrode for measuring the electron current. Immediately in front of the filament is a partition with a small axial hole, and a similar partition is in front of the collecting electrode. Electric fields parallel to one another and normal to the tube axis can be applied between two pairs of small plates, one pair of which is placed immediately behind the first partition and the other immediately in front of the second. The separation of these pairs of plates along the tube axis d is very large compared with the dimensions of the plates. If the filament is maintained at a potential $-V$ with respect to the partitions and an electric field of the form $E_0 \sin \omega t$ is applied between each pair of plates, show that the relationship between V and ω for which a current reaches the collector is $e/m = \omega^2 d^2 / 2\pi^2 V n^2$, where n is an integer.

2.11. The pole faces of an electromagnet are square, parallel plates of length l on each side, and their separation is small compared to l . The pole pieces are insulated from one another so that an electric field \mathbf{E} as well as a magnetic induction \mathbf{B} can be maintained between them. A beam of ions of charge e and mass M and of a variety of energies enters the region between the poles along a center line parallel to one edge. A photographic plate which is sensitive to these ions is placed at a distance d along from the far edge of the pole pieces normal to the initial direction of the ions. If l is much less than d or the radius of curvature of the ions in the magnetic induction \mathbf{B} between the poles, show that the trace of the ions on the plate is a parabola given by

$$y^2 = \frac{B^2 e}{Em} ld \left(1 + \frac{l}{2d} \right) x$$

where y is an axis perpendicular to \mathbf{E} and \mathbf{B} , and x is an axis parallel to these vectors. Show that the smallest value of x for any parabola if the beam contains various types of ions is $Edl(1 + l/2d)/2V$, where V is the ion energy in volts.

2.12. Consider the two-dimensional problem in which ions diverge from a source in a fan-shaped beam of small included angle, the central ray of which impinges normally upon one side of a wedge-shaped area in which there is a magnetic induction normal to the plane of the ion beam. The energy of the ion beam is such that the central ray emerges from the wedge-shaped area normal to the far edge. Show that to a first approximation the ion beam is brought to a focus at the point of intersection of the central ray with a line passing through the source and the apex of the wedge area.

2.13. A particle of charge e and mass m is projected circumferentially with a velocity v_0 between the plates of a cylindrical condenser in a plane normal to the axis. If the electric field at the point of projection is E_0 and there is also a uniform axial magnetic induction B , show that the particle will describe a circular orbit if

$$r_0 = \frac{mv_0^2}{e(E_0 + Bv_0)}$$

where r_0 is the radial distance of the point of projection from the axis. If the particle departs somewhat from the circular orbit, i.e., $r = r_0(1 + x)$, where x is small, show that it recrosses the circle of radius r_0 after traversing an angle $\phi = \pi(1 + eE_0r_0/mv_0^2)^{-\frac{1}{2}}$.

2.14. Ions are produced by electron impact in a small volume in a region of uniform magnetic induction B and are accelerated for a very short time normal to B by a potential pulse of duration t applied between two grids equally spaced on either side of the ion source. At the end of the period t the potential difference between the grids returns to zero and the ions move freely in a circular path around to the region in which they were formed, where they reach a collector at the end of a time interval T . The potential pulse of magnitude V is of such brief duration that the ions do not pass through either grid before the potential difference returns to zero. (a) Show that the radius of curvature of the ions ρ is equal to Et/B , where E is the field between the grids when the pulse is on. (b) Show that the quantity $eB^2\rho^2/mV$ must be less than unity. (c) Determine the ionic mass in atomic-weight units that would correspond for a singly charged ion to $T = 62 \mu\text{sec}$ if the magnetic induction is 0.1 weber m^{-2} .

2.15. Consider a small, rigid circular ring of radius r carrying a total charge e uniformly distributed around it which rotates about its axis-of-symmetry line through the center normal to the ring with the angular velocity ω in a region of uniform magnetic induction \mathbf{B} . Show that the force and torque on an element of the ring subtending an angle $d\phi$ at the center are

$$d\mathbf{F} = \frac{e}{2\pi} (\boldsymbol{\omega} \times \mathbf{r}) \times \mathbf{B} d\phi$$

$$d\mathbf{T} = \frac{e}{2\pi} (\boldsymbol{\omega} \times \mathbf{r})(\mathbf{r} \cdot \mathbf{B}) d\phi$$

Choosing $\boldsymbol{\omega}$ along one of the axes and expressing the other vectors in terms of their components and ϕ , integrate these expressions, and show that the force vanishes and the torque is given by $\frac{1}{2}er^2\boldsymbol{\omega} \times \mathbf{B}$.

2.16. If \mathbf{B} in the preceding problem is not uniform, show that there is a net force on the ring given by $\frac{1}{2}er^2\nabla(\boldsymbol{\omega} \cdot \mathbf{B})$.

2.17. Using the definition of generalized momentum $\mathbf{p} = m\dot{\mathbf{r}} + e\mathbf{A}$, and of magnetic moment $\mathbf{u} = (e/2)\mathbf{r} \times \mathbf{v}$, show from Eq. (2.21) that $\mu = -\partial H/\partial B$.

2.18. A beam of atoms each of which possesses a magnetic moment $e\hbar/2m_0$ enters the region between the parallel, rectangular poles of an electromagnet where the magnetic induction is B . The beam lies in the plane of symmetry between the poles and it initially makes an angle θ_1 with the normal to the edge at which it enters. If the energy of each ion is conserved and the moment lines up parallel or antiparallel to \mathbf{B} , show that the beam is diffracted and eventually makes an angle θ_2 with the normal, where $\sin \theta_2/\sin \theta_1 = \left(1 \pm \frac{\hbar e}{2m} \frac{B}{U}\right)^{\frac{1}{2}}$ and U is the kinetic energy of the atoms in the beam. How can this arrangement be used to polarize completely the entering beam of atoms?

2.19. It may be shown from the definition of the vector potential and the magnetic induction (in the preceding chapter) that the magnetic induction due to the motion of a charge e with a velocity \mathbf{v} at a distance r from it is

$$\mathbf{B}_{ev} = \frac{e}{4\pi\epsilon_0 c^2} \frac{\mathbf{v} \times \mathbf{r}}{r^3}$$

By applying this result to the rotation of a small, rigid circular ring of total charge e and area a at the rate ω about its axis of symmetry, which produces a magnetic moment $\mathbf{u} = (ea/2\pi)\omega$, it may be shown that the magnetic induction at a point distant r is given by $B_\mu = \nabla \times [\mathbf{u} \times \nabla(1/r)]/4\pi\epsilon_0 c^2$. Neglecting particular angles, show that the ratio of $|B_\mu|$ to $|B_{ev}|$ is of the order of $m_\mu \Delta r/e\hbar r$ and that in consequence the experiment in which the magnetic moment of an electron would be determined by a magnetometer measurement of the magnetic induction produced by it at a distant point is infeasible.

2.20. A uniform solid sphere of mass M carries a uniformly distributed surface charge e of electricity. It is set spinning about an axis through its center. Show that the ratio of its magnetic moment to its angular momentum is $5e/6M$.

2.21. A thin rod of length l has a uniform volume charge density. Show that its quadrupole moment is $l^2/6$. A thin disk of radius R has a uniform volume charge density. Show that its quadrupole moment is $-R^2/2$.

2.22. Assuming that an atomic nucleus has the form of a figure of revolution, a section through the axis of which is an ellipse, and that the nuclear charge is uniformly distributed throughout the volume, show that the quadrupole moment is $\frac{2}{5}(a^2 - b^2)$, where a is the elliptical semiaxis in the direction of symmetry and b is the equatorial radius. Taking the radius of the deuteron (H^2 nucleus) as 2×10^{-15} m and the quadrupole moment as 0.274×10^{-30} m 2 , show that the spheroid is prolate and that $a/b \cong 1.08$.

GENERAL REFERENCES

- Aston, F. W.: "Mass Spectra and Isotopes," Edward Arnold & Co., London, 1933.
 Bloembergen, N.: "Nuclear Magnetic Relaxation," Martinus Nijhoff, 1948.
 Born, M.: "Atomic Physics," chaps. 1, 2, 3, Blackie & Son, Ltd., Glasgow, 1946.
 Cajori, F.: "History of Physics," The Macmillan Company, New York, 1929.
 Compton, A. H., and S. K. Allison: "X-Rays in Theory and Experiment," chaps. 5, 9, D. Van Nostrand Company, Inc., New York, 1935.
 Crew, H.: "The Rise of Modern Physics," The Williams & Wilkins Company, Baltimore, 1928.
 Darrow, K. K.: *Bell System Tech. J.*, **32**, 74, 384 (1953).
 Finkelburg, W.: "Atomic Physics," chap. 2, McGraw-Hill Book Company, Inc., New York, 1950.
 Fraser, R.: "Molecular Rays," Cambridge University Press, London, 1931.

- Frisch, O. R., editor: "Progress in Nuclear Physics," vol. II, N. F. Verster, pt. I, Academic Press, Inc., New York, 1952.
- Harnwell, G. P.: "Principles of Electricity and Electromagnetism," 2d ed., chap. 10, McGraw-Hill Book Company, Inc., New York, 1949.
- Hughes, A. L., and L. A. Dubridge: "Photoelectric Phenomena," McGraw-Hill Book Company, Inc., New York, 1932.
- Magie, W. H.: "Source Book in Physics," McGraw-Hill Book Company, Inc., New York, 1935.
- Millikan, R. A.: "Electrons, Plus and Minus," University of Chicago Press, Chicago, 1935.
- Perrin, J.: "Atoms," Constable & Co., Ltd., London, 1923.
- Richtmyer, F. K., and E. H. Kennard: "Introduction to Modern Physics," chaps. 2, 3, 11, McGraw-Hill Book Company, Inc., New York, 1932.
- Segre, E., editor: "Experimental Nuclear Physics," vol. 1, N. F. Ramsey, pt. III, K. T. Bainbridge, pt. V, John Wiley & Sons, Inc., New York, 1953.
- Thomson, J. J., and G. P. Thomson: "Conduction of Electricity through Gases," Cambridge University Press, London, 1928.
- Van Vleck, J. H.: "Electric and Magnetic Susceptibilities," Oxford University Press, London, 1932.
- Zachariasen, W. H.: "X-ray Diffraction in Crystals," John Wiley & Sons, Inc., New York, 1945.

CHAPTER 3

ATOMIC STRUCTURE

3.1. Early Spectroscopy. Sir Isaac Newton, in 1666, was the first person to show that the dispersion of white light into a spectrum by a prism was an angular separation of colors originally present and not a production of these colors by the prism itself. If two similar prisms are so oriented that the first refracts the beam in one sense and the second refracts it in the other, the beam finally emerges uncolored. Newton did not use a sufficiently narrow slit to exhibit any line structure, and it was not until nearly a century later that Melville, in 1752, observed the existence of emission lines from an incandescent gas. Somewhat later, at the beginning of the nineteenth century, Wollaston and Fraunhofer observed the many black absorption lines in the solar spectrum. By 1850 the techniques of spectroscopy were sufficiently advanced for Bunsen and Kirchhoff to study the emission and absorption lines characteristic of incandescent gas sources. However, it was not until Rowland ruled good diffraction gratings in 1882 that sufficient resolution was available to exhibit the great number of very narrow lines characteristic of atomic spectra. It can be shown that the resolving power of a grating is equal to the diffraction order times the number of rulings that compose it. The use of gratings of 10,000 lines or more yielded a great quantity of precise data which it was recognized must be intimately related to atomic properties or characteristics.

The early efforts by the physicists of the past century to bring order out of this wealth of observation were conditioned by their background of classical mechanics and their experience with vibrating systems. It was natural to look for harmonic relationships of the form $\nu = n\nu_0$ for integral values of n since this was a familiar property of vibrating strings. However, this approach was unfruitful, and the first clue to a satisfactory interpretation was discovered in 1883 by Hartley, who observed recurring frequency differences. In 1885 Balmer achieved the following empirical representation of a series in the visible region of hydrogen:

$$\bar{\nu} = \frac{1}{\lambda} = R \left(\frac{1}{m^2} - \frac{1}{n^2} \right)$$

Here $\bar{\nu}$, called the wave number, is the reciprocal of the wavelength, which is the quantity primarily measured, and R is a constant, known as the *Rydberg constant*. Balmer's series was the one for which $m = 2$ and $n = 3, 4, 5$, etc. Later, Lyman observed the series for which $m = 1$, $n = 2, 3, 4$, etc., and Paschen the series for which $m = 3$, $n = 4, 5, 6$, etc. The Brackett and Pfund series are those for which m is 4 and 5, respectively. Closely analogous series, which could be represented by differences of fractions of the form $1/(m + p)^2$ where p was an approximate constant for the series, were found in the spectra of the alkali metals lithium, sodium, and potassium.

The principal simplifying step in the analysis of spectra was taken by Ritz in his recognition of the fact that the wave numbers of spectrum lines, whether forming members of obvious series or not, could be represented more simply as differences between characteristic wave numbers. These characteristic wave numbers are called *terms* (see Fig. 3.1). The Ritz combination principle represented a major simplification in the description of a spectrum. If $\bar{\nu}$'s represent the observed wave numbers and T 's represent terms, $T_a - T_b = \bar{\nu}_{ab}$.

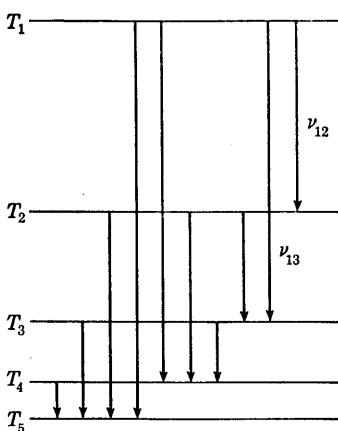


FIG. 3.1. Economy in the description of emitted frequencies as term differences.

The number of lines that can be represented in the extreme case of all possible combinations of n_T terms is the number of combinations of these in pairs, or $n_T(n_T - 1)/2$. This is greater by the factor $(n_T - 1)/2$ than the number of terms. Hence it is easily seen that 100 terms will

describe 5000 lines. The first step in the analysis of a complex spectrum is the search for common frequency differences between pairs of lines to determine term differences. For instance,

$$\begin{aligned}\bar{\nu}_{14} - \bar{\nu}_{24} &= (T_4 - T_1) - (T_4 - T_2) = (T_2 - T_1) \\ &= (T_3 - T_1) - (T_3 - T_2) = \bar{\nu}_{13} - \bar{\nu}_{23}\end{aligned}$$

A table of term differences can then be constructed, and the terms may be classified in various groups according to certain distinguishing characteristics exhibited by the spectrum lines associated with them. Thus some lines are sharp and others rather diffuse; some occur singly and others in close groups of doublets or triplets. The presence of electric or magnetic fields affects different groups of lines in different ways, splitting them into characteristic patterns. Also the presence of certain groups

of lines and the relative intensities of others depend on the method of exciting the spectrum, such as the use of an arc or a spark.

The term differences may be expressed as term values if some fiducial term is given an arbitrary value. When series are present, which can be represented by some empirical formula such as that of Balmer, a consistent scheme is achieved by assigning the term with the value zero to the limit. The families of terms can then be arranged in separate columns in such a way that the great majority of spectrum lines are represented by

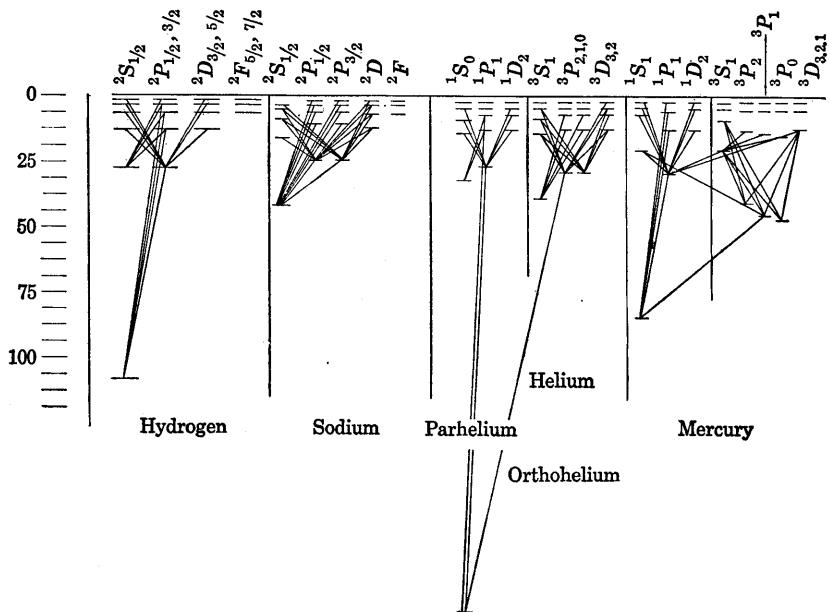


FIG. 3.2. Representative Grotrian diagrams. Transitions between terms in one ordinate interval, and those in the interval three above or three below it, correspond to spectrum lines in the visible. Transitions between closer terms are infrared lines, and those between more widely separated terms lie in the ultraviolet.

differences between the terms in neighboring columns. Such a diagram, displaying graphically the families of terms and indicating the predominant spectrum lines as transitions between them, is called a *Grotrian diagram* (Fig. 3.2). Labels *S*, *P*, *D*, etc., standing for certain early designations of the associated spectrum lines Sharp, Principal, Diffuse, etc., are given the term families, and superscripts indicate the degree of multiplicity of the close term levels. Instances are given in the accompanying diagrams. Easily distinguished series are characteristic of the groups of elements at the left of the periodic table, i.e., the alkali metals and alkaline earths. Series relations are much less obvious for the other elements. Morphological relations are evident among the series in a group of

elements, such as Li, Na, K, Rb, and Cs, and also between these series obtained in a low-voltage arc and certain series for Be, Mg, Cs, Sr, and Ba obtained in a high-voltage spark. The organization of these multifarious relationships provides the empirical structure which any successful theory must describe.

The major step in the interpretation of terms as elements in the description of atomic structure was taken by Bohr in 1913 by extending the Einstein energy-frequency relationship to term values. The Bohr frequency relationship states that the energy radiated by an atom in emitting a spectrum line of frequency ν is the difference between two characteristic energy levels represented by the initial and final term values:

$$h\nu = h\nu c = hc(T_i - T_f) = E_i - E_f \quad (3.1)$$

Thus the lowest term represents the normal state of the atom, in which it has no excess energy to radiate. The higher terms are excited states which represent a quantized excess energy content which can be disposed of by radiation when the atom returns to a lower-energy state. The structure of an atom is then characterized by certain specific and precisely defined energy states which Bohr associated with the position and motion of the constituent electrons. The terms in the ordinary Grotrian diagram represent states which essentially characterize the electronic structure for various conditions of binding of the most loosely held electron in the structure. If two or more electrons were to change their conditions simultaneously, additional terms would have to be used to describe the emitted radiation. The Grotrian type of diagram is the one most commonly encountered because it describes the most mobile or loosely bound of the atomic electrons, which most easily changes its condition under the influence of an external disturbance such as the impact of an electron or the absorption of a photon. Other diagrams of terms, such as X-ray energy-level diagrams, have to be used to describe the dependence of structure energy on the condition of an inner electron.

In the Bohr picture the term-series limit represents the removal of the outermost electron (the normal ionization process). The lowest term represents the normal condition of the atom in the lowest-energy, or ground, state. This latter is confirmed by observations on absorption spectra such as the Fraunhofer lines in the solar spectrum. The atoms of a gas, in going from the ground state to higher, excited states, are found to absorb from white radiation only those frequencies which represent so-called *allowed transitions* from the ground state. Allowed transitions are those for which emission lines are also observed. It is to be noted in the accompanying diagrams that there are certain low states from which return to the ground state by the emission of radiation

is not indicated. These transitions are not possible for reasons that will later become evident. Hence, if the atom is brought into such a state, it will remain there in a seminormal condition until a collision or some other event alters this condition. Such states are known as *metastable*, and they last an appreciable time. Therefore the absorption of radiation of an appropriate frequency to elevate the atom from such a state by an allowed transition to a higher state would be anticipated. Absorption spectra exhibit such events when metastable atoms are present.

Direct and independent verification of the Bohr interpretation of terms as energy levels was afforded by the experiments of Franck and Hertz in 1914. In these experiments electrons from a thermionic source were accelerated through a gas at low pressure. After suffering a number of atomic collisions, the electrons' energy was measured as a function of the potential that accelerated them initially. It was found that they lost little or no energy until their own energy exceeded a critical value which could be associated with an atomic transition from the ground state to the lowest excited state. Thus the interpretation of terms as energy states and transitions between them as energy conversions was directly confirmed. Characteristic energy losses representing transitions to higher excited states were observed, and, finally, additional free electrons were found to appear when the energy of the initial electrons exceeded the value corresponding to the difference between the ground level and the series limit on the Grotrian diagram, thus confirming the interpretation of this limit as the removal of the most loosely bound atomic electron. Other types of experiments on the absorption and subsequent emission of photons by atoms and molecules and the photoionization by radiation of sufficiently high frequency confirm the interpretation of atomic energy levels that has been outlined above.

3.2. Evidence for the Nuclear Atom. The early spectroscopic investigations yielded no structural information or suggestions as to the actual configurations represented by the association of the relatively massive positive charge and the lighter negative charge or charges in an atomic system. Justification for the assumption that the only mass associated with a negative charge is that of the electron comes from mass-spectrographic analyses of the charged particles produced in a gas discharge. Negatively charged ions are observed, but further experiments indicate that these are atoms to which an additional electron is temporarily attached. Ions that have as many as five or six negative charges removed, i.e., of net positive charge ne , are observed. Within the precision of the measurements, their charge-to-mass ratio is ne/M , where M is the mass of the neutral atom. The first evidence tending to provide a structural picture of the relationships between the elementary particles composing an atom came from experiments by Rutherford, Geiger, and

Marsden, 1906–9, on the scattering of α particles by thin foils.¹ α particles had been shown to be very energetic particles with the charge-to-mass ratio of a helium atom from which two electrons had been removed. They were observed to be emitted with energies in the range of several million electron volts from certain radioactive substances in the course of disintegration. The minimum uncertainty in our knowledge of the location of the α particle is related to its energy E by

$$\Delta x \cong \frac{\hbar}{\Delta p_x} = \frac{\hbar}{\sqrt{2ME}}$$

Assuming an energy of 5 Mev, corresponding to a doubly charged α particle which has fallen through an effective potential of 2.5×10^6 volts in leaving the nucleus emitting it, $\Delta x = 4 \times 10^{-15}$ m. This is very small

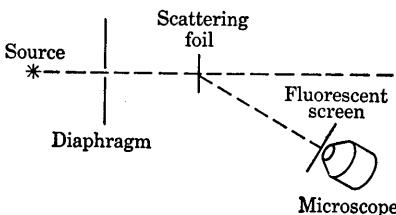
in comparison with atomic dimensions, which are of the order of 10^{-10} m. Hence α particles are essentially infinitesimal test particles, and α scattering experiments can be analyzed in a simple, classical manner without reference to any wave properties.

FIG. 3.3. Observation of α -particle scattering.

to approach sufficiently close to the center of the scattering atom to probe its structure. The classical distance of closest approach is essentially the distance r_{\min} at which the potential energy of the particle in the Coulomb field of the scattering atom equals the initial kinetic energy of the α particle: $E = 2Ze^2/r_{\min}$, so $r_{\min} = 2Ze^2/E = \frac{1}{2} \times 10^{-13}$ m for 5-Mev α particles on lead. Hence, these radioactive α particles should be satisfactory probes of atomic structure.

Geiger and Marsden studied the number of α particles as a function of angular deflection ϕ from their initial trajectory as a result of passing through a thin metal foil. The α particles were detected by observing the scintillations produced when they impinged on a fluorescent screen as in Fig. 3.3. Chadwick, in a refinement of the experiment, in which an annular ring of foil was used to increase the solid angle, made still more precise measurements of the angular deviations. Geiger and Marsden used silver and gold foils only a few atomic diameters thick, and Chadwick used platinum foils as well. The α -particle energy could be varied by placing thin screens of retarding material over the source, and the foil thickness also could be varied. It was found that the number of

¹ E. Rutherford, J. Chadwick, and C. D. Ellis, "Radiations from Radioactive Substances," Cambridge University Press, London, 1930.



scintillations per second per unit area of the fluorescent screen was proportional to (1) $\sin^{-4}(\phi/2)$, (2) the foil thickness t for small thickness, (3) the square of the reciprocal energy of the α particles, and (4) the square of the serial number of the scattering element in the atomic sequence, that is, its atomic number. The atomic number was identified with the total positive charge on the nucleus of the atom by these experiments and by those of Barkla¹ and Moseley² on X-ray spectra.

The analysis of scattering experiments depends critically upon the assumptions that are made regarding the distribution of charge and mass in an atom. A proposal made by Thomson was that the massive positive charge was distributed uniformly over a considerable portion of the volume occupied by the atomic system and that it was penetrable by the α particle. The negatively charged atomic electrons with their relatively small masses were assumed to be readily displaceable and without any considerable influence in deflecting the α particle. The important point about an encounter with such an atom is that in no case is the particle deviated from its path through any large angle. If the α particle passes directly through the center of the atom, it can be seen by symmetry considerations that there is no lateral displacement. The maximum displacement occurs when the trajectory just grazes the boundary of the positive-charge distribution. If the particle passes a distance b from the center of charge, a short calculation yields its small angular displacement. For the great majority of the interactions with the atoms in the foil, the distance of closest approach between the α particle and the atom is very large compared to the dimensions of the positive-charge distribution. Hence the number of collisions in which the particle actually passes through this charge may be neglected as a complication which has but slight influence on the results, and the deflecting charge may be considered to be concentrated at a point. Referring to Fig. 3.4, which represents the plane of the trajectory,

$$F_x = \frac{dp_x}{dt} = \frac{Zee'}{4\pi\epsilon_0} \frac{1}{(z^2 + b^2)} \frac{b}{(z^2 + b^2)^{\frac{3}{2}}}$$

¹ C. G. Barkla and C. A. Sadler, *Phil. Mag.*, **16**, 550 (1908).

² H. G. J. Moseley, *Phil. Mag.*, **27**, 703 (1914).

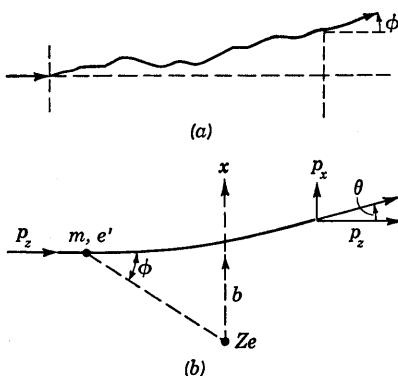


FIG. 3.4. (a) Multiple encounters in a foil projected on a plane containing the beam axis. (b) Single-encounter small-angle Coulomb scattering (exaggerated).

Assuming $v_z = dz/dt$ is effectively constant, the integration may be expressed in terms of the angle ϕ as

$$P_x = \frac{Zee'}{4\pi\epsilon_0 v_z} \int_0^\pi \frac{1}{b} \sin \phi \, d\phi$$

or

$$\frac{P_x}{P_z} = \tan \theta \cong \theta = \frac{Zee'}{2\pi\epsilon_0 p_z v_z} \frac{1}{b}$$

For nonrelativistic velocities the angular deflection is thus equal to the ratio of the potential energy of the particles at the distance of closest approach to the kinetic energy of the α particle.

If the radius of the positive charge is even as small as a hundredth the separation of atoms in crystals and if Z is of the order of 10, the maximum potential energy is of the order of 10^4 ev. As α -particle energies are of the order of 10^6 ev, θ is very small for the greatest possible single deflection.

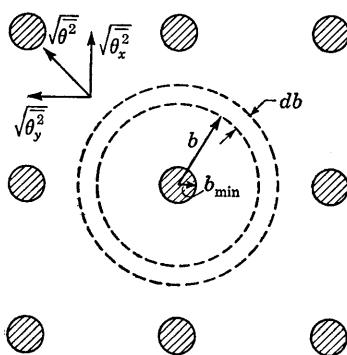


FIG. 3.5. Scattering by an atomic charge distribution in a crystal.

The average value of θ after the many encounters made in traversing a foil is zero, but the average value of θ^2 in the plane of the orbit is finite and it is the most convenient parameter for carrying out the subsequent analysis. The limits of integration in calculating the average are taken as b_{\min} and b_{\max} , and it is assumed that b_{\max} is sufficiently greater than b_{\min} that the square of the latter can be neglected in the denominator. Using the geometry of Fig 3.5,

$$\begin{aligned} \bar{\theta}^2 &= \left(\frac{Zee'}{2\pi\epsilon_0 p_z v_z} \right)^2 \frac{\int_{b_{\min}}^{b_{\max}} 2\pi b (1/b^2) \, db}{\int_{b_{\min}}^{b_{\max}} 2\pi b \, db} \\ &= 2 \left(\frac{Zee'}{2\pi\epsilon_0 p_z v_z b_{\max}} \right)^2 \ln \frac{b_{\max}}{b_{\min}} \end{aligned}$$

The argument of the logarithm depends on the particular assumptions made regarding the size of the atomic charge distribution, but because of the logarithmic dependence the result is not very sensitive to this choice. If b_{\max} is taken of the order of the interatomic spacing of a crystal and b_{\min} about a hundredth of this value, the logarithm factor is 4.6. As the deflection is small at each individual scattering event, the total angular deviation in passing through a foil is, on this theory, the sum of the deflections for all the collisions made. These deflections

will take place in all planes that pass through the axis of the beam of particles, and to sum the results of all collisions the components along each of two mutually perpendicular axes at right angles to the beam must be combined separately. If x and y are such axes, the squares of the momentum components along them due to a collision for which the trajectory lies in a plane making an angle ψ with the x axis are proportional to $\cos^2 \psi$ and $\sin^2 \psi$, respectively. The average values of these factors for all possible angles ψ are each $\frac{1}{2}$; thus the root-mean-square average deflection along each of these axes at a collision is the square root of $\bar{\theta}^2/2$. Writing this as $\bar{\theta}_{\text{rms}}$,

$$\bar{\theta}_{\text{rms}} = \sqrt{\bar{\theta}_y^2} = \sqrt{\bar{\theta}_x^2} = \frac{Zee'}{2\pi\epsilon_0 p_z v_z b_{\max}} \sqrt{\ln \frac{b_{\max}}{b_{\min}}}$$

The probability of observing a particular deflection, say $\phi_x = m\bar{\theta}_{\text{rms}}$, after n encounters can be calculated as in Chap. 6. The number N_ϕ of α particles scattered at the angle ϕ_x after n collisions is

$$N_\phi = N_0 e^{-\phi_x^2/n\bar{\theta}_x^2} = N_0 e^{-m^2/n}$$

where $n = \pi N' b_{\max}^2 t$ if N' is the number of atoms per unit volume and t is the foil thickness.

The important feature of this probability formula is its exponential dependence on angle. This does not agree with the observed $\sin^{-4}(\phi/2)$ angle dependence. Geiger had shown that the mean angular deflection for α particles passing through a 0.5×10^{-6} -m foil of gold was about 1° . Hence the likelihood of observing scattering at large angles would be vanishingly small. Geiger actually observed more than 1 in 8000 α particles to be deflected through an angle greater than 90° by a similar foil. Hence the theory proposed by Thomson, in which the scattering results from many small deviations, is definitely in disagreement with experiment.

Rutherford Scattering. In order to account for the experimental results on α -particle scattering, it was then proposed by Rutherford that the events leading to the large angular deflections of these particles on passing through a foil are primarily single scattering events resulting from close encounters between the α particles and highly concentrated, positively charged mass centers of the same order of magnitude in spatial extent as the α particle itself. On this concept the α particle does not penetrate the region of positive atomic charge but is scattered by it simply as by a point mass M having a charge Ze . As the force of interaction is assumed to be electrostatic, the variation of the force with distance is of the inverse-square type and the problem is essentially the same as calculating the deflection suffered by a comet in passing near the sun. This calcu-

lation will be found in detail in many places, and it is merely outlined in the following paragraph.

The simplest starting point is the expression for the energy of the two particles in plane polar coordinates, since the motion is confined to the plane defined by the line of centers of the particles and the initial relative velocity. The nonrelativistic case only will be considered, as illustrated in Fig. 3.6.

$$\text{Potential energy} = \frac{2Ze^2}{4\pi\epsilon_0 r}$$

$$\text{Kinetic energy} = \frac{1}{2}M_\alpha[\dot{r}_1^2 + (r_1\dot{\eta}_1)^2] + \frac{1}{2}M[\dot{r}_2^2 + (r_2\dot{\eta}_2)^2]$$

where r is the separation of the particles and η is the azimuthal angle of the line joining them. The expression is simplified by transforming to the center of gravity system in which by definition $r_1M_\alpha = r_2M$, $r = r_1 + r_2$, and $\dot{\eta}_1 = \dot{\eta}_2 = \dot{\eta}$. The total energy may then be written in terms of r , η , $M_r = MM_\alpha/(M + M_\alpha)$, and the momenta $p_r = M_r\dot{r}$ and $p_\eta = M_r r^2 \dot{\eta}$ as

$$H = \frac{1}{2M_r} \left[p_r^2 + \left(\frac{p_\eta}{r} \right)^2 \right] + \frac{2Ze^2}{4\pi\epsilon_0 r}$$

The Hamiltonian equations

$$\dot{p}_r = -\frac{\partial H}{\partial r} \quad \dot{p}_\eta = -\frac{\partial H}{\partial \eta}$$

$$\dot{r} = \frac{\partial H}{\partial p_r} \quad \dot{\eta} = \frac{\partial H}{\partial p_\eta}$$

then yield

$$\ddot{r} = \left(\frac{p_\eta}{M_r} \right)^2 r^{-3} + \frac{2Ze^2}{4\pi\epsilon_0 M_r} r^{-2}$$

and

$$p_\eta = M_r r^2 \dot{\eta} = \text{constant}$$

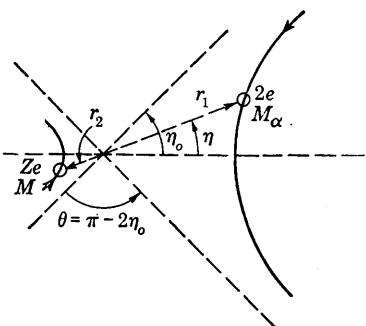
FIG. 3.6. Scattering of an α particle by a nucleus of charge Ze and mass M , center of mass at rest.

These may be simplified by writing $u = 1/r$ and considering r as a function of η , which is in turn a function of t . Then

$$\frac{\partial^2 u}{\partial \eta^2} + u = -B \quad \text{where } B = \frac{2Ze^2 M_r}{4\pi\epsilon_0 p_\eta^2}$$

This is the simple-harmonic differential equation, and the solution is $u = B(\cos \eta/\cos \eta_0 - 1)$, where the constants of integration are so chosen that $\eta \rightarrow \eta_0$ for great values of r , and u is a maximum or r is a minimum for $\eta = 0$.

To calculate the scattering of α particles by a foil it is convenient to express the angular deviation $\theta = \pi - 2\eta_0$ as a function of the distance b at which the two particles would pass if there were no interaction. This



latter quantity is the limiting value of $r \sin(\eta_0 - \eta)$ as r becomes very large. Writing $\eta_0 - \eta = \delta$, the orbit equation yields

$$\frac{1}{r} = B \frac{\cos(\eta_0 - \delta) - \cos \eta_0}{\cos \eta_0} = B(\cos \delta - 1) + Bb \frac{\tan \eta_0}{r}$$

or as r becomes very large $\cos \delta \rightarrow 1$ and $b \rightarrow (\cot \eta_0)/B$. B can also be expressed in terms of b^2 through $p_\eta = M_r v_\alpha b$, where v_α is the α -particle velocity at a great distance. Eliminating B and p_η ,

$$\begin{aligned} b &= \frac{2Ze^2 M_r}{4\pi\epsilon_0 (M_r v_\alpha)^2} \tan \eta_0 = \frac{2Ze^2}{4\pi\epsilon_0 M_r v_\alpha^2} \cot \frac{\theta}{2} \\ &= \frac{1}{2} r_{\min} \cot \frac{\theta}{2} \end{aligned}$$

The distance r_{\min} is the closest possible approach for a head-on collision, as may be seen by equating the initial kinetic energy of the particles when at a great distance from one another to their mutual potential energy when at rest at the distance of closest approach.

$$\frac{1}{2} M_r v_\alpha^2 = \frac{2Ze^2}{4\pi\epsilon_0 r_{\min}}$$

Using this expression for b , the number of α particles scattered in the angle between θ and $\theta + d\theta$ is the number passing a nucleus in the radial range from b to $b + db$. The probability of a particle passing a scattering center within this radial range is proportional to $2\pi b db$ and, if G particles per unit area per unit time impinge on a foil of thickness t containing N' scattering centers per unit volume, the number experiencing encounters in the range b to $b + db$ per unit area per unit time is

$$n_b db = 2\pi b N' t G db$$

Using the relation between b and θ ,

$$n_\theta d\theta = \frac{G}{8} 2\pi N' t r_{\min}^2 \cot \frac{\theta}{2} \csc^2 \frac{\theta}{2} d\theta$$

and the fraction of the particles scattered per unit solid angle is this quantity divided by $2G\pi \sin \theta d\theta$. Thus

$$\begin{aligned} N' t \sigma(\theta) &= \frac{N' t}{16} r_{\min}^2 \csc^4 \frac{\theta}{2} \\ \sigma(\theta) &= \left(\frac{r_{\min}}{4} \right)^2 \csc^4 \frac{\theta}{2} \end{aligned} \quad (3.2)$$

$\sigma(\theta)$ is the cross section for scattering at the angle θ in the system in which the center of gravity is at rest. This theory is seen to describe

properly the experimental dependence of scattering on θ and t , on the nuclear charge Ze , and also on the α -particle energy as required by the observations of Geiger and Marsden.

Actually, of course, the scattering process is observed in the laboratory system of coordinates for which the nucleus is initially at rest and the center of gravity is in motion. If the mass of the scattering nucleus is not much greater than the mass of the α particle, very important corrections may be involved. However, for an aluminum scatterer the correction involved in the transformation to the laboratory system is only a few per cent, and with heavy elements it is quite negligible. A careful series of experiments by Chadwick using foils of copper, silver, and platinum yielded values for r_{\min} from which the following values of Z were deduced: Cu = $29.3e$, Ag = $46.3e$, Pt = $77.4e$. The correct values are $29e$, $47e$, and $78e$, respectively; hence it can be seen that this technique not only verifies the correctness of the atomic model proposed by Rutherford but also determines the actual charge borne by the scattering nucleus to within about 1 per cent. In conclusion it should be mentioned that the inverse-square scattering law breaks down for values of r_{\min} of the order of 10^{-14} m. This is of the order of magnitude of nuclear dimensions, and additional forces which are as yet very imperfectly understood become of importance at these very small distances.

Confirmation of the fundamental significance of the sequential number of the elements in the order from the lightest, hydrogen, through the periodic table to uranium, the heaviest of the permanent elements, came from the observations of Moseley in 1913. He made a systematic photographic study of the X-ray spectra emitted by a series of targets from sodium through zinc, identifying two groups of radiations by their different absorption coefficients in aluminum. The more penetrating of these is known as K radiation, and the less penetrating is known as L radiation, a designation previously given by Barkla. The crucial observation by Moseley was that if the square root of the frequency of either one of these types of radiation is plotted against the serial number of the element with which it is associated the points lie very accurately on a straight line. In accordance with the planetary theory of atomic structure propounded by Bohr and Rutherford, this serial number, which is called the *atomic number* Z , was identified with the net nuclear positive charge Ze . These observations and later extensions of them cleared up uncertainties as to the correct sequential order (as in the case of argon and chlorine) and verified certain gaps in the periodic table representing elements that had not been observed up to that time.

The result of the α -particle scattering experiments and the observations by Moseley was the establishment as a component of the atom of a high concentration of positive charge representing the major amount of

the atomic mass, the charge on this central nucleus in units of e being equal to the serial number of the element in the atomic sequence. These observations yielded no information on the configuration of the electron structure relative to this nucleus. Indeed it is evident from the uncertainty relations that it is not permissible to specify with precision the spatial disposition of electrons about a nucleus. Writing $\Delta x = \hbar/\sqrt{2mE}$ from the uncertainty relation and putting into this expression the values representing an electron with an energy of 100 ev, it is found that Δx is of the order of 1.7×10^{-11} m. Atomic dimensions are themselves of the order of 10^{-11} to 10^{-10} m. Thus with the known orders of magnitude of energies of electrons in atoms, the spatial localization that is permissible is no less than the atomic dimensions themselves. Of course, this does not mean that there is no structure in the permissible sense of well-defined energy levels resulting from the association of electrons with the atomic nucleus. It is not possible, however, to specify the positions of the electrons giving rise to this structure, although electron-density or probability calculations are possible.

The general theory of multiple small-angle scattering developed in considering the Thomson atom hypothesis can be applied to the scattering of fast electrons such as the β particles emitted in radioactive processes with energies of the same order of magnitude as those of the α particles. The primary difference between the α - and β -particle cases is that for the same kinetic energy the associated wavelength is much greater for the lighter particle, and the classical description in terms of point masses is inadequate for the small distances of approach that lead to large-angle scattering. The multiple-small-angle-scattering theory is found to account for the observed results. Very-high-energy electron scattering experiments yield little specific information on the electron structure of atoms, but the scattering of less-energetic electrons and of radiation does provide useful information, as is seen in Sec. 5.1.

The identity in pattern of the X-ray emission lines of the elements clearly indicates close similarity and even qualitative identity of energy-level structure among the elements for those electrons having a high binding energy which are associated with X-ray emission and absorption. The energy level or term diagram of atomic X-ray levels is approximately the same except for a scale factor for all X-ray spectrum patterns. Also the optical-spectrum-pattern similarities between the arc spectrum (essentially nonionized) of an element and the spectrum of the n th-ionized atom n elements further along the atomic sequence confirm the structural similarities of extranuclear electrons in atomic systems. Something is known of the spectrum of copper from which as many as 19 electrons have been removed. Again something is known of all the spectra of phosphorus for which from 0 to 13 electrons have been removed.

The knowledge of the highly excited spectra in this region near the center of the periodic table is very extensive, and the analogies between successive excited spectra of increasing order are very pronounced. These spectra for which the atomic number differs but the number of electrons is the same are known as *isoelectronic sequences*, and their patterns confirm the similarities in structure which are characteristic of a given number of electrons $Z - n$ surrounding a nucleus of positive charge Ze .¹

Additional evidence of similar import comes from the mass-spectrographic analysis of the products of ionization in a gas by electron impact or photoelectric absorption. Since the mass spectrograph measures e/M , it may be used to determine the initial occurrence of the successive stages of ionization for which the net charge is $1e$, $2e$, $3e$, etc., as the energy of the electron producing the ionization is increased. In this way the first five ionization potentials of mercury have been measured, and they agree as they should with the series limit of the successively ionized mercury atoms as determined from spectroscopic evidence. The mass-spectrographic method outlined above is a particularly illuminating one for studying the energetics of atomic and molecular systems through the dependence of the number and nature of the products of ionization on the energy of the ionizing electrons and the other circumstances relevant to the process.²

3.3. Experimental Determination of Energy Levels. By far the most precise information available on the structure of atomic systems comes from measurement of the frequencies or wavelengths of the electromagnetic radiation that is emitted or absorbed by these systems. These measurements can be made with very great precision indeed, as is evidenced, for instance, by the narrowness of spectrum lines. The width of such lines, however, is generally considerably greater than the inherent breadth produced by the uncertainty in atomic energy levels. One reason for line broadening is the instrumental limitation imposed by lens and mirror aberrations, diffraction, etc. A second is that the atomic or molecular systems are in motion with respect to the observer, and if this motion has a random component along the line of sight there is a spread in the observed frequency. The first-order Doppler-effect frequency shift is $\Delta\nu = (\nu_0 v \cos \theta)/c$ where v is the atomic velocity and θ is the angle between v and the direction of observation. From the velocity distribution law (Sec. 6.2) the number of atoms with a velocity between v_x and $v_x + dv_x$ is proportional to $e^{-Mv_x^2/2kT}$. Therefore, taking the x axis as the direction of observation, the intensity distribution as a

¹ H. E. White, "Introduction to Atomic Spectra," chap. 17, McGraw-Hill Book Company, Inc., New York, 1934; A. G. Shenstone, *Repts. Progr. in Phys.*, **5**, 210 (1938).

² H. D. Smyth, *Rev. Mod. Phys.*, **3**, 347 (1931); A. G. Gaydon, "Dissociation Energies," Chapman & Hall, Ltd., London, 1947.

function of the frequency is $e^{-\frac{Mc^2}{2kT} \left(\frac{\nu_0 - \nu}{\nu_0}\right)^2}$, and the frequency difference due to the Doppler effect between points on the curve which are half the intensity at $\nu = \nu_0$ is

$$\delta\nu_d = \frac{2\nu_0}{c} \sqrt{(\ln 2) \frac{2kT}{M}} = 1.67 \frac{\nu_0}{c} \sqrt{\frac{2kT}{M}}$$

The limitation in precision of measurement imposed by the Doppler effect is then

$$\frac{\delta\nu_d}{\nu_0} = \frac{\delta\lambda_d}{\lambda_0} = 1.67 \sqrt{\frac{2kT}{Mc^2}} \quad (3.3)$$

Thus it is least for low temperatures and large masses, but even for hydrogen at room temperature, for which $Mc^2 \approx 10^9$ ev and $kT \approx 2.5 \times 10^{-2}$ ev, the precision is 1 part in 10^5 . The line is of course narrower if there is less variation of velocity along the line of sight, and for this reason transverse observation of atomic or molecular beams is used to further decrease line broadening from this cause. A third reason for line broadening is that external influences disturb the atomic system under observation. Electric and magnetic fields from neighboring atoms shift atomic energy levels, and also atomic collisions interfere with the radiation process.

Collision broadening results essentially from a limitation of the time during which coherent radiation can be emitted by the atomic system. The average time between atomic collisions in a gas can be calculated from the average atomic velocity and the average distance traversed between collisions. The mean velocity can be found from the energy equipartition law (Sec. 6.3), $\frac{1}{2}Mv^2 = \frac{3}{2}kT$. The mean free path l is roughly the reciprocal of the number of atoms contained in a cylinder of radius equal to an atomic diameter and unit length, that is,

$$l = (N'\pi d^2)^{-1}$$

where d is the effective diameter for a collision sufficiently close to affect the radiation process and N' is the number of atoms per unit volume. The time between collisions is then

$$t_c = \frac{l}{v} = \frac{1}{N'\pi d^2} \sqrt{\frac{M}{3kT}}$$

By Eq. (2.19),

$$2\pi \delta\nu_c = \frac{1}{t_c}$$

and

$$\frac{\delta\nu_c}{\nu_0} = \frac{\delta\lambda_c}{\lambda_0} = \frac{1}{2}N'd^2\lambda_0 \sqrt{\frac{3kT}{Mc^2}} \quad (3.4)$$

This is about 0.4 times the number of atoms in a volume $d^2\lambda_0$ greater than the Doppler broadening. Its magnitude turns out to be very nearly equal to the Doppler broadening for atmospheric pressure and for wavelengths in the visible region of the spectrum.

However, spectrum lines would not be of zero width even if all the above causes of broadening could be made negligible. This inherent breadth comes from the finite length of time involved in the radiation process. The value of this inherent breadth properly calculated from quantum-mechanical principles is not different in order of magnitude from that implied by the classical discussion leading to Eq. (1.26). Here the half width of a line on a frequency scale was seen to be

$$\begin{aligned}\delta\omega &= \frac{2\omega_0^2 r_0}{3c} \\ \frac{\delta\nu_0}{\nu_0} &= \frac{\delta\lambda_0}{\lambda_0} = \frac{4\pi}{3} \frac{r_0}{\lambda_0} \cong \frac{12 \times 10^{-15}}{\lambda_0}\end{aligned} \quad (3.5)$$

where λ_0 is in meters. Thus, for the visible spectrum, $\lambda_0 \cong 6 \times 10^{-7}$ m and $\delta\lambda_0/\lambda_0 = 2 \times 10^{-8}$, which is about 10^{-2} times the order of magnitude of the analogous quantities in Eqs. (3.3) and (3.4).

Low-energy Atomic Transitions. The techniques for investigating the relatively-low-energy atomic transitions below, say, 10^{-3} ev, or wavelengths greater than 1.24 mm, are quite different from those useful in the higher energy range. Certain of the methods appropriate to this region were mentioned in the preceding chapter, though that discussion related primarily to characteristic frequencies of precession of atomic nuclei in magnetic fields. The molecular-beam method of Rabi and his colleagues is one of the most powerful for these investigations.¹ Its general nature is similar to that of the Stern-Gerlach experiment, but the innovation of focusing and analyzing by inhomogeneous magnetic fields renders it a method of precision. A beam of neutral atoms or molecules issues from a narrow slit into a long, highly evacuated region. It first encounters an inhomogeneous magnetic induction normal to the beam, in which particles with magnetic moment μ are deflected up or down depending on the orientation of the magnetic moment. Of those having a slight upward component of velocity on leaving the entrance slit, the ones with properly oriented moments will be deflected downward and approxi-

¹ See J. B. M. Kellogg and S. Millman, *Rev. Mod. Phys.*, **18**, 323 (1946) and D. R. Hamilton, *Am. J. Phys.*, **9**, 319 (1941).

mately focused on the central slit, and the others will be lost from the beam. Analogously, those starting downward with oppositely directed moments will be deflected upward and approximately focused on the central slit, the others being lost. Of those which are so focused, the slowly moving particles will graze the poles and rapidly moving ones will remain near the axis. In the next inhomogeneous field the opposite sense and magnitude of deflection will occur as in Fig. 3.7 owing to the oppositely directed gradient of the magnetic induction, and the beam will pass through the slit in front of the detector. However, should something occur to change the orientation of an atomic moment in the region indicated by the homogeneous magnetic induction \mathbf{B} , the force

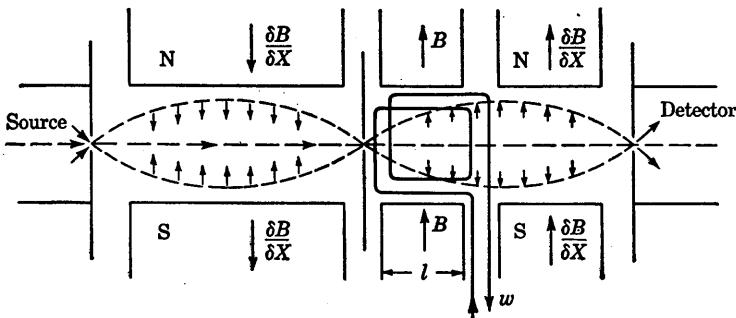


FIG. 3.7. Schematic depiction of the traversal of an atomic-beam apparatus by a beam of atoms having magnetic moments.

on such an atom would not be such as to tend to return it to the axis and it would not be focused on the detector slit but would be lost to the poles in the second inhomogeneous field. This decrease in beam strength would be noted by the detector, and thus the device provides a sensitive method of detecting any change in the projection of the magnetic moments along the x axis, normal to the beam of particles.

The importance of the molecular-beam technique depends upon the fact that it provides a sensitive method for detecting a change in the \mathbf{B} component or a reorientation relative to \mathbf{B} of the magnetic moments of the atoms composing the beam. A perturbing high-frequency electromagnetic field of variable angular frequency ω at right angles to the homogeneous magnetic induction \mathbf{B} is brought about by a circuit element such as a long, narrow coil in that neighborhood. It may be shown that resonance of ω with the precessional frequency of the magnetic moments in \mathbf{B} brings about such a reorientation.¹ The resonant condition is observed by a decrease in intensity of the detected beam. Thus γ is measured in terms of the frequency of ω by Eq. (2.22): $\gamma = \mu/l = -\omega/B$. If \mathbf{B} is perfectly homogeneous, the resonant frequency may be determined

¹ I. I. Rabi, *Phys. Rev.*, **51**, 652 (1937).

to a high degree of precision by making the time of comparison, determined by the velocity of the beam and the length of the region of uniform \mathbf{B} , long in relation to the period of precession. It is not feasible to make \mathbf{B} sufficiently uniform over a great enough length to take full advantage of the potential precision, but a modification recently proposed by Ramsey¹ overcomes this difficulty. In his arrangement, shown in Fig. 3.8, the perturbing field is produced by two coils over shorter regions near the ends of the beam path between the pole pieces over each of which individually the magnetic induction is sufficiently uniform

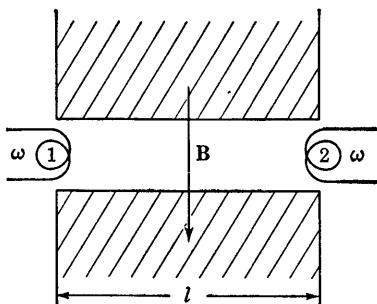


FIG. 3.8. Ramsey's arrangement of two perturbing coils to improve the precision of atomic-beam experiments.

to permit the theoretical resolution to be approached. Though the path length near each coil is considerably shorter than l of the figure, the relative phase of ω and the perturbation it produces in the moment orientations is significant so that, when the perturbation caused by the second coil is added to that produced at an earlier time (l/v) by the first coil, full advantage is taken of the time of flight of the atoms over l with the velocity v and the requirement of uniformity of \mathbf{B} between the coils is relaxed. The situation is analogous to that of the interference pattern produced by two slits in a screen, with widths equal to the coil dimensions and separated by a distance l , for light of angular frequency ω incident on the screen; the structure of the interference pattern is much finer than the diffraction patterns of the individual slits. The products of \hbar and the resonant ω 's measured by the atomic-beam technique represent energy differences between atomic states in the prevailing \mathbf{B} . As ω can be compared very precisely by means of harmonic techniques with frequency standards broadcast from national laboratories, which are known to 1 part in 10^7 , very accurate values of the energy differences can be obtained. It should be noted that, in distinction to the constant characteristic γ 's of elementary particles, the γ for an atomic or molecular system is in general a function of \mathbf{B} because of the dependence of the mean relative orientations of the components of the atomic moment on the external magnetic induction (Sects. 4.3 and 4.4). By extrapolation to zero \mathbf{B} the normal atomic-energy-level differences can, however, be obtained.

The other general technique for use in this region ($\lambda > 1.24 \times 10^{-3} \text{ m}$) is that of microwaves.² This war-born technique is a very powerful

¹ N. F. Ramsey, *Phys. Rev.*, **76**, 996 (1949).

² W. E. Lamb, *Repts. Progr. in Phys.*, **14**, 19 (1951); B. Bleany, *Repts. Progr. in*

one, and in common with the molecular-beam method it is generally used in absorption rather than emission. However, measurements in emission have been made by Dicke¹ and his collaborators on water vapor in the atmosphere and by Ewen and Purcell² on the radiation of the hyperfine-structure frequency of hydrogen in the plane of the galaxy. Though Cleeton and Williams³ made the first measurements of this type, the general utilization of microwaves awaited the development of efficient high-frequency generators such as the klystron and the resonant-cavity

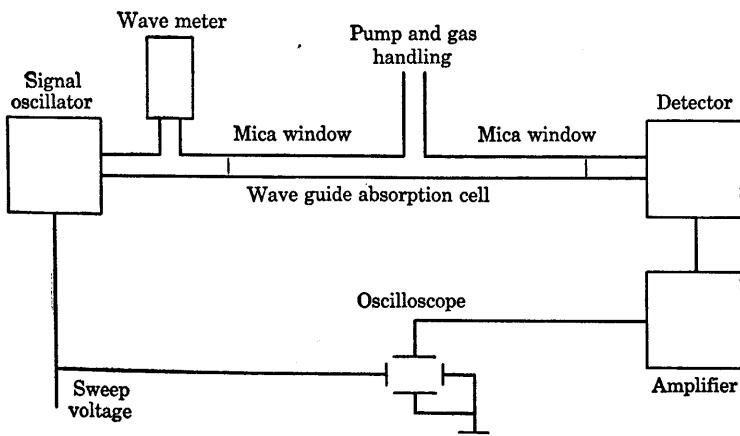


FIG. 3.9. Schematic of simple microwave spectrometer.

magnetron. By the use of standard klystrons, and harmonic generation by a nonlinear element such as a crystal rectifier, wavelengths from 2 or 3 mm on up through the centimeter region can be generated. These waves are transmitted through rectangular wave guides of dimensions properly chosen for the frequency band to be used. Wavelengths greater than twice the longer dimension of the guide are rapidly attenuated. As indicated in Fig. 3.9, regions of the guide or special cavities segregated by transparent windows are used to contain the sample whose absorption is being studied. Detection, or a determination of the power level, is generally accomplished by a heterodyne type of detector. Attenuators and directional couplers are used to reduce undesired interaction between circuit elements.

The frequency can be measured by comparison with signals broadcast by national standards laboratories or by local measurements using

Phys., **11**, 178 (1947); W. Gordy, *Revs. Mod. Phys.*, **20**, 668 (1948); Y. Ting and D. Williams, *Phys. Rev.*, **82**, 507 (1951).

¹ R. H. Dicke, B. Beringer, R. L. Kyhl, and A. B. Vane, *Phys. Rev.*, **70**, 340 (1946).

² H. I. Ewen and E. M. Purcell, *Phys. Rev.*, **83**, 881 (1951).

³ C. E. Cleeton and N. H. Williams, *Phys. Rev.*, **45**, 234 (1934).

the resonance of a cavity of known dimensions. Also, the klystron frequency can be modulated to sweep through the frequency range to be investigated and the modulating frequency used as a time axis on an oscilloscope to display the absorption curve as a function of klystron frequency. Many different experimental techniques have been developed by the groups at Columbia, Harvard, MIT, Oxford, and North Carolina. These include the application of alternating electric fields and magnetic fields to the sample to produce atomic-energy-level perturbations and consequent modulated absorption. As can be seen from Eq. (3.4), pressure broadening is particularly important in this region of long wavelengths, but the actual absorption is small and hence it is not feasible to work at extremely low pressures. Very large cavities or very long wave-guide sections must frequently be used to observe absorption peaks at all. Finally, the fact that $\hbar\nu$ is much less than the mean thermal energy kT introduces complications not present at high energies. By the Boltzmann theorem (Secs. 6.2 and 6.5) the ratio of the population of an upper and lower energy state separated by $\hbar\omega$ is $e^{-\hbar\omega/kT}$, or, as $\hbar\omega \ll kT$, the relative difference in population is approximately $\hbar\omega/kT$. This fraction is so small that absorption is almost canceled by induced emission, and the net energy absorption is proportional to ω^2 . This fact leads to rapidly decreasing absorption coefficients with lower frequencies, and measurements become infeasible with present techniques at wavelengths greater than about 10 cm.

In addition to determining the frequencies of maximum absorption, which yield the molecular energy-level differences, the magnitudes of the absorption coefficients which yield the probabilities of the molecular transitions can be obtained by these methods. Quantitative measurements can be made using wave-guide absorption cells as indicated in Fig. 3.9 or by measuring the amplitude of oscillation sustained in a resonant cavity containing the absorbing material when the cavity is supplied with electromagnetic energy at a constant rate. The power level in a cavity, P , measured by a crystal detector connected to the cavity is proportional to the square of the resonant Q value of the cavity. If the Q values of the cavity are measured with and without the absorbing substance, as Q_1 and Q_0 , respectively, the attenuation may be calculated. By definition

$$Q = - \frac{\omega U}{\partial U / \partial t}$$

where ω is the resonant angular frequency and U is the electromagnetic energy in the cavity. The attenuation coefficient of a substance traversed by a wave is by definition

$$\alpha = - \frac{1}{U} \frac{dU}{dx} = - \frac{1}{U} \frac{dU}{dt} \left(\frac{dx}{dt} \right)^{-1} = - \frac{1}{v_p U} \frac{dU}{dt}$$

and, as $v_p = \omega\lambda/2\pi$, $\alpha = 2\pi/\lambda Q$, and the attenuation due to the absorbing gas in the cavity is

$$\alpha = \frac{2\pi}{\lambda} \left(\frac{1}{Q_1} - \frac{1}{Q_0} \right)$$

Higher-energy Atomic Transitions. The techniques for investigating large atomic-energy intervals are quite different from those discussed above and they have a much longer history.¹ They are associated with the field of classical optics and differ at almost every point from microwave techniques but most fundamentally in that the ultimate standard of comparison is a length rather than a time interval. The primary wavelength standard is a well-defined spectrum line in the red region of the cadmium spectrum. This has been very accurately compared by interferometric methods with the standard meter, in terms of which its wavelength is $6438.4696 \times 10^{-10}$ m under certain standard conditions with a precision of about 1 part in 10^7 . Wavelengths in the infrared are customarily expressed in microns, μ , where 1μ is 10^{-6} m, and wavelengths in the visible and ultraviolet are generally given in angstroms A, where 1 A is 10^{-10} m. Three or four hundred secondary-standard lines from about 2500 to 7000 A have been established by comparison with the primary standard to about the same precision and are used to determine regional wavelength scales. It is clear that measurements in the optical and microwave regions will be concordant only if the copulative datum, the velocity of light, is known to the same accuracy. At present this is not the case. The nominal precision in c has until 1949 been thought to be about 130 parts in 10^7 , but inconsistencies observed in geodetic surveys² cast some doubt on this. Recent work indicates that the old value of c was in error by a still larger amount. One of the accurate measurements of c has been made by K. D. Froome,³ who has used a microwave interferometer to determine the free-space wavelength of electromagnetic radiation whose frequency could be compared with a standard quartz-clock frequency. A Pound stabilized klystron oscillator provided microwaves with a frequency of 24,005 Mc, corresponding to a wavelength of about 1.25 cm. These oscillations are compared with a 5-Mc quartz oscillator whose frequency was scaled up to 250 Mc by direct frequency multiplication and thence to 24,000 Mc by harmonic conversion. The 5-Mc oscillator was compared directly with the National Physical Laboratory standard quartz clock. The wavelength was measured by detecting the interference minimum for two different

¹ R. A. Sawyer, "Experimental Spectroscopy," Prentice-Hall, Inc., New York, 1944; G. R. Harrison, R. C. Lord, and J. R. Loofbourow, "Practical Spectroscopy," Prentice-Hall, Inc., New York, 1948.

² C. I. Aslakson, *Trans. Am. Geophys. Union*, **30**, 475 (1949); Summary of Recent Determinations of c , J. F. Mulligan, *Am. J. Phys.*, **20**, 165 (1952).

³ K. D. Froome, *Proc. Roy. Soc. (London)*, **213A**, 123 (1952).

positions of the mirror in the variable arm of the interferometer. This mirror displacement, 161.74 cm, was equivalent to 259 minima and was measured with gauge blocks whose length was compared both with the Hg¹⁹⁸ standard wavelength and with the British National metre. The experimental arrangement is shown in Fig. 3.10. The junction at A acts as a beam splitter sending radiation into the shorted "fixed" line to the left and the "variable" arm to the right. The reflected waves join again at A and interfere in the detector arm. The small amount of

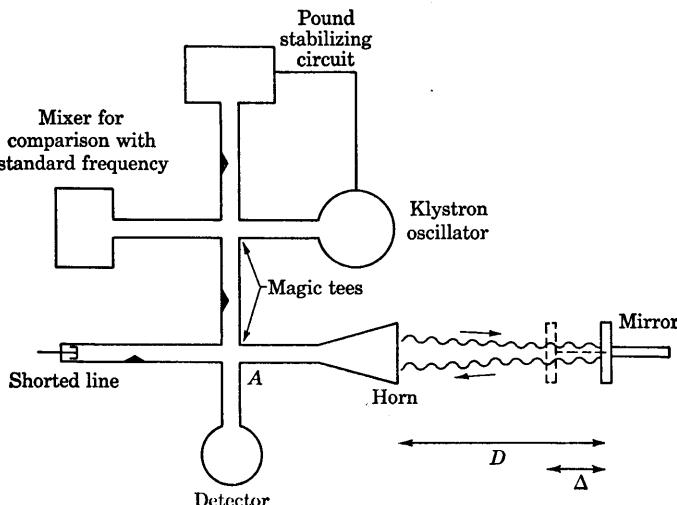


FIG. 3.10. Microwave interferometer. (After Froome.)

reflected energy (10^4 of radiated energy) necessitates a superheterodyne detector which uses a 24,050-Mc auxiliary oscillator to give an intermediate frequency of 45 Mc, which can be amplified. The mirrors used ranged from 15×15 cm to 30×30 cm, and the distance D was varied from 6.5 to 21.5 m. These variations allowed a check in the corrections which were applied for the diffraction effects introduced by the large ratio of wavelength to mirror and horn size. The accuracy of a single setting to a minimum was about $\pm 3 \mu$. Numerous measurements were made at different horn-mirror distances, with different mirror sizes and with different observers, and the average value of c after corrections was $299,792.6 \pm 0.7 \text{ km sec}^{-1}$.

E. Bergstrand¹ has used an amplitude-modulated beam of visible light (5600 Å) transmitted over a 6-km base line to determine c . As shown in Fig. 3.11 a light beam modulated by a Kerr cell is reflected by a mirror back to a phototube detector. The phototube sensitivity is also varied with the modulated frequency so that a resonance effect

¹ E. Bergstrand, *Arkiv Fysik*, **3**, 479 (1950).

occurs for the modulation waves. As the mirror is moved, the measured intensity goes through minima corresponding to quarter wavelengths of the modulation waves. The product of the modulation wavelength and the modulation frequency gives the velocity of light. A 50-cycle square wave was used to reverse the phase of the modulation and to gate the detector alternately into a difference amplifier, thus increasing the sensitivity of detecting zero points as the frequency was varied. The difference between the mirror positions for several hundred zero points was measured to be 5413.370 m when the modulation frequency was 8,332,387 cycles per second. The result of repeated measurements gave 299793.1 ± 0.2 km sec $^{-1}$ for the velocity of light. DuMond and Cohen¹ give 299792.9 ± 0.8 km sec $^{-1}$ for their least-squares adjusted value, a

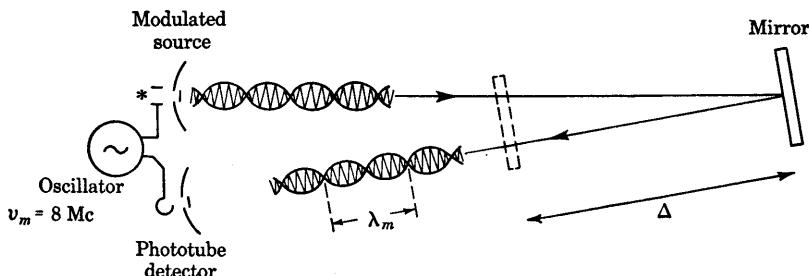


FIG. 3.11. Velocity-of-light apparatus. (After Bergstrand.)

precision of about 27 in 10^7 . It is becoming increasingly important for atomic physics that the value of this fundamental constant be determined with an accuracy of about 1 part in 10^7 .

The optical region, for our purposes, may be defined roughly as extending from $\lambda = 1.24 \times 10^{-8}$ m, or 10^{-3} ev, to the wavelength or energy of the radiation necessary to liberate the most tightly bound electron from the heaviest atom, uranium, which is $\lambda = 10^{-11}$ m, or 1.24×10^5 ev. At still higher energies in the gamma- or cosmic-ray regions the observational techniques again differ quite fundamentally from those of use at lower energies. In the optical region both emission and absorption methods are employed. In emission the sources are the atoms themselves excited by combustion or photon or electron impact in electric arcs, sparks, or X-ray tubes. In absorption, suitable continuous sources, such as incandescent solids, arcs, or white X radiation (*Bremsstrahlung*), are used as a background against which to observe absorption. Again, detection devices vary considerably over the optical region. The thermocouple provides a method useful in the entire region, and it is almost the only known detecting device for the far infrared where the wavelength is greater than 10^{-5} m. It is not a particularly sensitive device since its

¹ J. W. M. DuMond and E. R. Cohen, *Rev. Mod. Phys.*, **25**, 691 (1953).

minimum detectable power is of the order of 10^{-11} watt, equivalent to more than 5×10^8 quanta per second in the far infrared. Of course, it becomes much more efficient as a detector of quanta at higher frequencies. Photocells and photoconductive cells are used for the region from 10^{-5} to nearly 10^{-7} m and are from 10^2 to 10^5 times as sensitive as the thermocouple in this region. In the visible region the photomultiplier cell has a detection threshold of about 10^{-16} watt, corresponding to a minimum of about 100 quanta per second. The photographic plate can be used for all wavelengths less than about 10^{-6} m, and by reason of its cumulative effect it is capable of detecting the incidence of very few quanta over an area comparable in size to a developable emulsion grain. It has a region of least usefulness for wavelengths between 10^{-7} and 10^{-9} m, which is the far ultraviolet or soft X-ray region. In the X-ray region, $\lambda \leq 10^{-8}$ m, the photoionization produced by the passage of photons through a gas and the fluorescence produced by photons in crystals also provide detection techniques which have the advantage of rapidity and greater ease of quantitative interpretation than the photographic plate.¹

The common techniques which tend to unify the optical category are those by which the photon energies are distinguished from one another, i.e., dispersion or diffraction. The former is limited to those regions where the absorption of dispersive substances is not excessive and where the change of index of refraction with wavelength is large enough to be useful. In these regions between the infrared and the ultraviolet, lenses can be used to increase intensity and to form images. The resolution that can be obtained with a prism is limited by the diffraction pattern produced by the aperture determined by the prism and is defined as this aperture times the rate of change of angular deviation with wavelength. Alternatively, the resolving power R may be written $R = t dn/d\lambda$, where t is the length of the prism base and n is the index of refraction. The pertinent constants of some materials that are used for refractive and dispersive elements are listed as follows:

TABLE 3.1

Material	Region of transparency, m $\times 10^{-6}$	n_{vis}	$(d\theta/d\lambda)_{vis}$, A $^{-1} \times 10^{-5}$
Fluorite	0.12 - 9	1.434	0.333
Quartz	0.185 - 3.5	1.544	0.628
Rock salt	0.02 - 17	1.544	0.938
Glass	0.36 - 2	1.6	1

¹ Ultraviolet techniques: J. C. Boyce, *Rev. Mod. Phys.*, **13**, 1 (1941); W. C. Price, *Repts. Progr. in Phys.*, **14**, 1 (1951). Infrared techniques: Van Z. Williams, *Rev. Sci. Instr.*, **19**, 135 (1948).

Thus materials of this type can be used over a range of about 100-fold in wavelength, and the resolving power for an opening of the order of 10 cm is about 10^4 . This is seen to be considerably smaller than microwave resolving powers. The dispersion is a function of the wavelength, and empirical expressions must be used for interpolation between secondary-standard lines which are photographed on the same plate for reference. Higher resolution in the neighborhood of the visible region can be obtained by the use of interferometric elements such as etalons or Lummer plates placed in the spectrograph train yielding a high order of local dispersion. In this way the fine structure of a spectrum line can be easily investigated with a resolving power of the order of 5×10^5 .

Diffraction is available for wavelength measurement over a very much greater wavelength range than dispersion. When used with lenses, diffraction techniques are of course limited to the ranges of the preceding paragraph. However, focusing mirrors or curved diffraction gratings may be used to extend the range of usefulness from the far infrared through the X-ray region. The resolving power of a diffraction grating is simply equal to the number of lines it contains times the order of diffraction being used. With present techniques, resolving powers of 5×10^5 have been achieved, but in general use the resolving power is less by a factor of about 10. There are also certain disadvantages accompanying the use of a grating. The overlapping of orders of diffraction complicates interpretation, but usually ambiguity can be avoided by the use of filters which limit the wavelength range of radiation admitted to the instrument. Another disadvantage of a grating instrument is the astigmatism due to the use of off-axis spherical mirrors. In the prism spectrograph the lines on the photographic plate are stigmatic images of the slit; in the grating instrument they are the superposition of astigmatic images of the points constituting the slit line.

It is most convenient to have the grating serve as its own focusing element, thus reducing the spectrograph to: entrance slit, diffracting and focusing element, and exit slit. Gratings for the long-wavelength optical region are ruled on concave spherical mirrors. The astigmatic focus condition is fulfilled if the source s and the focus f in Fig. 3.12 lie on a circle (Rowland circle) of diameter equal to the radius of curvature of the mirror grating and tangent to it. This can easily be seen by setting up the condition that rays following the paths $r + r'$ and $r_0 + r'_0$ of Fig. 3.12 arrive at f in phase.

$$r^2 = (r_0 \sin \alpha - R \sin \delta\theta)^2 + [r_0 \cos \alpha - R(1 - \cos \delta\theta)]^2$$

$$r'^2 = (r'_0 \sin \alpha' + R \sin \delta\theta)^2 + [r'_0 \cos \alpha' - R(1 - \cos \delta\theta)]^2$$

Expanding in powers of the assumed small quantity $\delta\theta$ and setting $R \delta\theta$ equal to d , the grating space in the plane of the figure, the condition is

$$n\lambda = d(\sin \alpha' - \sin \alpha) + \frac{1}{2}d^2 \left(\frac{\cos^2 \alpha}{r_0} - \frac{\cos \alpha}{R} + \frac{\cos^2 \alpha'}{r'_0} - \frac{\cos \alpha'}{R} \right)$$

The first term yields the diffraction-grating condition and the second bracket vanishes making the curved-grating result the same as that for the plane grating if $1/R = (\cos \alpha)/r_0 = (\cos \alpha')/r'_0$, which is the condition that s and f lie on the Rowland circle. In the X-ray region, normal reflectivity is low but thin crystals curved in one dimension (in the plane of Fig. 3.12) can be used by taking advantage of the negative index of refraction for crystals in the X-ray range. Total reflection from the crystal takes place at grazing angles, and, at the expense of the greater

aberrations and smaller resolving power at such large angles with the normal, very high dispersion can be obtained (Sec. 2.2).

As indicated in the preceding discussion the precision with which the wavelength of spectrum lines can be determined is in certain cases as high as 1 part in 10^6 but more commonly of the order of 1 part in 10^5 . Term values, being differences between these reciprocal wavelengths, are known somewhat less precisely. The differences between close terms, however, involve the loss of a number of significant figures, and the

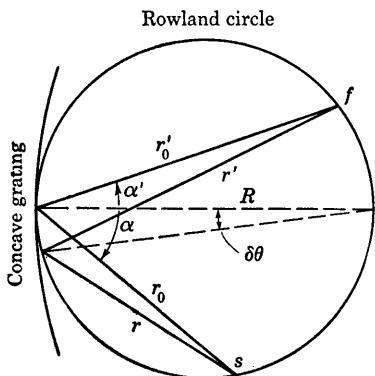


FIG. 3.12. Construction for deriving the grating focusing condition.

precision of fine-structure information may become very low indeed, even to merely ascertaining the order of magnitude. These differences, however, are the ones measured directly with high precision by microwave techniques. In the short-wavelength region, collision broadening becomes less important and that due to the Doppler effect and external fields predominates. In general the change in energy-level values caused by external electric fields (Stark effect) is more important than that caused by magnetic fields (Zeeman effect). The very broad spectrum lines are generally associated with electrons so related to the atom as to be particularly affected by neighboring electric fields or occupying energy states of unusually short lifetimes. One cause for short-lived states is the possibility, when the energy relations permit, of the energy of excitation of one electron being transferred to a less energetically bound electron enabling it to leave the atom as an ion. This is a nonradiative transition and may lead to a much shorter life and hence greater line breadth than that predicted by Eq. (3.5). Such a process is called *autoionization* or the *Auger effect*. In the hard X-ray region, where Stark, collision, and

Doppler broadening tend to become less important, the inherent radiation breadth is approached, and many lines are of the order of only 10^{-3} Å or 10^{-13} m in breadth.

3.4. Schroedinger Theory. The problem of atomic theory is how to formulate a description of an aggregate of two or more elementary particles, each of which has the properties outlined in Chap. 2, in such a way that an account is given of the known physical properties of such an aggregate; in particular that a description is given of the characteristic energy levels which are known in such detail and with such high precision. A single particle can be described by Eq. (2.17):

$$\psi = \psi_0 e^{-i(Et - \mathbf{p} \cdot \mathbf{r})/\hbar}$$

The next-simplest case is that of two particles which are distinguishable from one another. An example of such a simple system is the Bohr-Rutherford picture of a hydrogen atom, consisting of a proton characterized by its mass and electric and magnetic moments and an electron characterized by its mass and electric and magnetic moments. The distinguishing characteristic of this system as opposed to its two separate constituents is the energy of interaction between them. This suggests setting up a differential equation for which functions of the form of Eq. (2.17) will be solutions in the absence of this interaction and which will be in accord with the conservation laws of Chap. 1. A successful way to set about this, proposed by Schroedinger¹ in 1926, is to start from Eq. (1.5) as an expression for the conservation of energy and write \mathbf{p} and E as differentials of ψ , as implied in Eq. (2.17).

$$\frac{\partial \psi}{\partial t} = -\frac{i}{\hbar} E \psi \quad \text{and} \quad \frac{\partial \psi}{\partial x} = \frac{i}{\hbar} p_x \psi$$

or formally

$$E = i\hbar \frac{\partial}{\partial t} \quad \text{and} \quad \mathbf{p} = -i\hbar \nabla \quad \text{or} \quad p_x = -i\hbar \frac{\partial}{\partial x}$$

where these differentials operate on the function ψ . Then the Hamiltonian equation (1.5) becomes

$$H \left(-i\hbar \frac{\partial}{\partial q_i}, q_i \right) \psi = i\hbar \frac{\partial \psi}{\partial t} \quad (3.6)$$

which is known as *Schroedinger's equation*. Sans-serif type is used to distinguish a quantity which implicitly contains a differential operator.

Simple as the imaginative step leading to this equation looks, it constituted one of the major theoretical advances in atomic physics. Historically, the concepts involved in the establishment of the quantum

¹ E. Schroedinger, *Ann. Physik*, **79**, 361, 489 (1926); **81**, 109 (1926).

theory of atomic structure were due to Heisenberg.¹ However, the equivalent formalism which he employed follows somewhat less naturally than the presentation here adopted. A very large part of the success of our present description of atomic phenomena is due to the adequacy with which the values of ψ which satisfy Eq. (3.6) account for the facts of observation. The linearity in E which is equivalent to $i\hbar \partial/\partial t$ on the right of Eq. (3.6) is an essential characteristic; the relativistic equation (1.16) is not a suitable starting point because of the ambiguity of the resulting radical. The present discussion is limited to the nonrelativistic case. The physical interpretation of ψ is the same as that proposed in

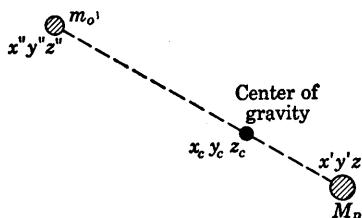


FIG. 3.13. Cartesian coordinates for the two-body problem.

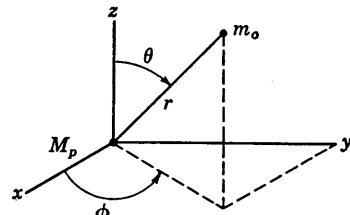


FIG. 3.14. Variables of separation for the hydrogen problem.

Chap. 2. The square of the absolute magnitude of ψ , which is in general complex, is proportional to the probability of observing the system described by it with the particular values of the coordinates of which ψ is a function. If ψ is complex and ψ^* its complex conjugate, $\psi\psi^* = |\psi|^2$, and $\int\psi\psi^* dv$ is proportional to the probability of observing the system within the range of the variables over which the integration is performed. The statement that the system exists within the total range of the variables, $\int\psi\psi^* dv = 1$, is a so-called *normalizing condition* that determines the constants of integration. Finally, the average or expectation value of any quantity f characteristic of the system of electron and proton is by a consistent definition $\bar{f} = \int\psi f\psi^* dv$.

Writing the electron mass m_0 and the proton mass M_p , the kinetic energy is

$$\frac{\frac{p_p^2}{M_p} + \frac{p_e^2}{2m_0}}{2M_p} = -\frac{\hbar^2}{2M_p} \left(\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} + \frac{\partial^2}{\partial z'^2} \right) - \frac{\hbar^2}{2m_0} \left(\frac{\partial^2}{\partial x''^2} + \frac{\partial^2}{\partial y''^2} + \frac{\partial^2}{\partial z''^2} \right) = -\frac{\hbar^2}{2M_p} \nabla'^2 - \frac{\hbar^2}{2m_0} \nabla''^2$$

where the primed coordinates specify the position of the proton and the double-primed ones the position of the electron, as in Fig. 3.13. The equation as a whole is simpler in terms of the coordinates (x_c, y_c, z_c) of the

¹ W. Heisenberg, *Z. Physik*, **33**, 879 (1925).

center of gravity of the two masses and their coordinates with respect to the center of gravity, $-m_0x/(m_0 + M_p)$, etc., and $M_p x/(m_0 + M_p)$, etc., where x , y , and z are the coordinates of m_0 with respect to M_p . Making the transformation to the variables of Fig. 3.14,

$$\begin{aligned} x_c &= \frac{M_p x' + m_0 x''}{M_p + m_0}, & y_c &= \frac{M_p y' + m_0 y''}{M_p + m_0}, & z_c &= \frac{M_p z' + m_0 z''}{M_p + m_0} \\ x &= x'' - x', & y &= y'' - y', & z &= z'' - z' \end{aligned}$$

The kinetic energy in terms of the total mass M and reduced mass m_r then becomes

$$-\frac{\hbar^2}{2} \left(\frac{1}{M} \nabla_c^2 + \frac{1}{m_r} \nabla^2 \right)$$

where $M = M_p + m_0$ and $1/m_r = 1/M_p + 1/m_0$. Writing V for the potential energy, Eq. (3.6) becomes

$$\left[\frac{\hbar^2}{2} \left(\frac{1}{M} \nabla_c^2 + \frac{1}{m_r} \nabla^2 \right) - V \right] \psi = -i\hbar \frac{\partial \psi}{\partial t}$$

It is assumed that ψ is separable in terms of the spatial and temporal variables, that is, $\psi = U(x, x_c, \dots) f(t)$. After this substitution is made, the equation is divided by ψ giving

$$\frac{1}{U} \left[\frac{\hbar^2}{2} \left(\frac{1}{M} \nabla_c^2 + \frac{1}{m_r} \nabla^2 \right) - V \right] U = -\frac{i\hbar}{f} \frac{\partial f}{\partial t}$$

The two sides involve independent variables, and for the equation to be true for all values of them it must be equal to a constant. Let $\frac{i\hbar}{f} \frac{\partial f}{\partial t} = E'$. A solution for f is $f = A_m e^{-(i/\hbar)E_m t}$. If, further, there is no force tending to move the center of gravity, V is not a function of x_c , y_c , and z_c . The equation for U may then be further separated in terms of

$$U = u_c(x_c, y_c, z_c) u(x, y, z)$$

into the two following equations, where the constant E'_m is written as the sum of two constants E_c and E_n referring to the motion of the center of gravity and the motion relative to it, respectively.

$$\frac{\hbar^2}{2M} \nabla_c^2 u_c = -E_c u_c \quad (3.7)$$

$$\frac{\hbar^2}{2m_r} \nabla^2 u - Vu = -E_n u \quad (3.8)$$

These equations are of the form $Hu = Eu$, in which the time does not appear. It is the differential equation giving the steady state of the system with the energy E . H is called the *Hamiltonian operator* or

energy operator, and the functions u which describe the physical state of the system are called the state functions or *energy eigenfunctions*. This is the typical form of the equations of atomic or quantum mechanics with which the subsequent discussion will largely deal. The energy values E for which the equation possesses acceptable solutions are said to be *energy eigenvalues* or observable energy values equivalent to the energy operator H . A consistent interpretation of Eq. (3.7) is that it represents the motion of the center of gravity. A solution is seen to be $u_c = A_c e^{\pm i \mathbf{p}_c \cdot \mathbf{r} / \hbar}$ where A_c is a constant and \mathbf{p}_c is a vector of magnitude $\sqrt{2M\bar{E}_c}$. Thus the translatory motion of the center of gravity is represented by a simple periodic function of the form of Eq. (2.17):

$$\psi_c = f_c u_c = \psi_{c0} e^{-i(E_{ct} \pm \mathbf{p}_c \cdot \mathbf{r})/\hbar}$$

The effective mass of the combined particles located at the center of gravity is $M = M_p + m_0$, and the momentum corresponds to the classical value $\sqrt{2M\bar{E}_c} = Mv_c$.

The motion of particular interest here, however, is the relative motion of M_p and m_0 , given by Eq. (3.8). To consider this further, V must be expressed specifically in terms of the coordinates determining it. The potential energy is made up of the two parts representing the interaction of the electric and magnetic moments involved. It can be seen by an order-of-magnitude calculation that the electrostatic energy is considerably the greater. As mentioned in Sec. 1.7 the magnetic induction due to a magnetic moment \mathbf{u}_1 at a distance R large compared to its dimensions is

$$\mathbf{B} = \frac{1}{4\pi\epsilon_0 c^2} \nabla \times \left(\frac{\mathbf{u}_1 \times \mathbf{R}}{R^3} \right)$$

The energy of a second moment \mathbf{u}_2 in the presence of the first is $-\mathbf{B} \cdot \mathbf{u}_2$, and the maximum value of the magnetic energy occurs when \mathbf{u}_1 and \mathbf{u}_2 are both perpendicular to the line joining them and is

$$E_m = \frac{2\mu_1\mu_2}{4\pi\epsilon_0 c^2 R^3}$$

On the other hand the electrostatic energy of two charges e and $-e$ a distance R apart is $E_e = -e^2/4\pi\epsilon_0 R$. Using the moments as those characteristic of electron motion, $\mu = e\hbar/2m_0$, the ratio becomes

$$\frac{E_m}{E_e} = \frac{\hbar^2}{2c^2 R^2 m_0^2} = \frac{1}{8\pi^2} \left(\frac{\hbar}{cm_0} \right)^2 \frac{1}{R^2} = \frac{1}{8\pi^2} \left(\frac{\lambda_c}{R} \right)^2$$

where λ_c is the Compton wavelength, which is about 2×10^{-12} m. Thus for atomic distances $R \approx 10^{-10}$ m, the ratio is of the order of 10^{-5} and the electrostatic energy is seen to exceed the magnetic energy by this factor.

The above considerations relate of course to the interactions between moments associated with the electron. It has been seen that the magnetic moment of a proton is less by a factor of the order of 10^3 than that of the electron, so that the electron-proton potential energy due to magnetic interactions is smaller than the electrostatic interaction by a factor of the order of 10^8 . Thus to the first approximation V may be taken as the electrostatic energy alone.

Taking the central charge as Ze for generality, Eq. (3.8) then becomes

$$\nabla^2 u + \frac{2m_r}{\hbar^2} \left(E_n + \frac{Ze^2}{4\pi\epsilon_0 r} \right) u = 0 \quad (3.9)$$

and the problem is to determine a solution u that will be acceptable as a physical description. The fact that the only variable that appears in V is the distance r suggests that polar coordinates are appropriate for the problem. In terms of these,

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

A solution can be found by assuming that u can be written as the product of three separate functions, each of which contains but a single one of the variables, that is, $u(r, \theta, \phi) = R(r)\Theta(\theta)\Phi(\phi)$. Substituting this in Eq. (3.9) and multiplying through by $(r^2 \sin^2 \theta)/u$ to display the term in ϕ , one obtains

$$\begin{aligned} \frac{1}{R} \sin^2 \theta \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + \frac{\sin \theta}{\Theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right) + \frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2} \\ + \frac{2m_r}{\hbar^2} r^2 \sin^2 \theta \left(E_n + \frac{Ze^2}{4\pi\epsilon_0 r} \right) = 0 \quad (3.9') \end{aligned}$$

The next-to-last term involves the independent variable ϕ alone, and for the equation to be true for all values of the independent variables this term must be equal to a constant. By common convention this constant is chosen as $-m^2$, i.e.,

$$\frac{d^2 \Phi}{d\phi^2} + m^2 \Phi = 0$$

Putting $-m^2$ for this term in Eq. (3.9') and multiplying the resulting equation by $1/\sin^2 \theta$ to exhibit the term in θ , one obtains

$$\begin{aligned} \frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + \frac{1}{\Theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right) - \frac{m^2}{\sin^2 \theta} \\ + \frac{2m_r r^2}{\hbar^2} \left(E_n + \frac{Ze^2}{4\pi\epsilon_0 r} \right) = 0 \end{aligned}$$

The second and third terms involve θ alone, and the first and fourth involve r alone; hence they must be equal to constants of equal magnitude but opposite sign. By convention the constant is chosen as $l(l + 1)$, and the three total differential equations into which Eq. (3.9) divides are

$$\frac{d^2\Phi}{d\phi^2} + m^2\Phi = 0 \quad (3.10)$$

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + \left[l(l + 1) - \frac{m^2}{\sin^2 \theta} \right] \Theta = 0 \quad (3.11)$$

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \left[\frac{2m_r}{\hbar^2} \left(E_n + \frac{Ze^2}{4\pi\epsilon_0 r} \right) - \frac{l(l + 1)}{r^2} \right] R = 0 \quad (3.12)$$

Angular-momentum Quantization. To be acceptable physically the solutions of these equations must be finite and single-valued, the various derivatives must exist, and if the system is to be localized in space the integral of uu^* over the range of the variables must be finite. For solutions of Eq. (3.10) to be single-valued, that is, $\Phi(\phi) = \Phi(\phi + 2\pi)$, m must be a positive or negative integer, and a solution is immediately seen to be

$$\Phi = A_\phi e^{im\phi}$$

where A_ϕ is a constant. Similarly, Eq. (3.11) is a well-known one, and physically acceptable solutions are customarily written

$$\Theta = A_\theta P_l^m(\cos \theta)$$

where A_θ is a constant and $P_l^m(\cos \theta)$ are the *associated Legendre functions*, which are simple polynomials in $\cos \theta$ defined by

$$P_l^m(\cos \theta) = \frac{(\sin \theta)^{|m|}}{2^l l!} \left[\frac{d}{d(\cos \theta)} \right]^{l+|m|} (\cos^2 \theta - 1)^l$$

From the definition, these polynomials exist only if l is an integer and equal to or greater than $|m|$. The angular equations are of such frequent occurrence as components in various problems that they warrant brief discussion. The product of Φ and Θ is known as a *tesseral harmonic* and is generally written $Y_{l,m}$. The constant $A_\phi A_\theta$ is determined by normalizing over a unit sphere, i.e., multiplying $Y_{l,m} Y_{l,m}^*$ by $\sin \theta d\theta d\phi$ and setting the integral over $0 < \theta < \pi$ and $0 < \phi < 2\pi$ equal to unity. In this way the first few harmonics may be shown to be

$$\begin{aligned}
 Y_{0,0} &= (4\pi)^{-\frac{1}{2}} & Y_{2,2} &= \left(\frac{15}{32\pi}\right)^{\frac{1}{2}} \sin^2 \theta e^{2i\phi} \\
 Y_{1,1} &= \left(\frac{3}{8\pi}\right)^{\frac{1}{2}} \sin \theta e^{i\phi} & Y_{2,1} &= \left(\frac{15}{8\pi}\right)^{\frac{1}{2}} \sin \theta \cos \theta e^{i\phi} \\
 Y_{1,0} &= \left(\frac{3}{4\pi}\right)^{\frac{1}{2}} \cos \theta & Y_{2,0} &= \left(\frac{5}{16\pi}\right)^{\frac{1}{2}} (3 \cos^2 \theta - 1) \\
 Y_{1,-1} &= \left(\frac{3}{8\pi}\right)^{\frac{1}{2}} \sin \theta e^{-i\phi} & Y_{2,-1} &= \left(\frac{15}{8\pi}\right)^{\frac{1}{2}} \sin \theta \cos \theta e^{-i\phi} \\
 & & Y_{2,-2} &= \left(\frac{15}{32\pi}\right)^{\frac{1}{2}} \sin^2 \theta e^{-2i\phi}
 \end{aligned}$$

These functions determine the angular dependence of the solution of Eq. (3.9) and in themselves are of considerable interest. $Y_{0,0}$ is seen to be a constant, and by forming $\sum_{m=l}^{-l} Y_{l,m} Y_{l,m}^*$ this is seen to be a constant also for any l . Thus though the Y 's are functions of the angles, $Y_{0,0} Y_{0,0}^*$ is spherically symmetric and the sums over m for the other real products $Y_{l,m} Y_{l,m}^*$ for any l are also spherically symmetric. Also the Y 's fall into two classes on the basis of their behavior at points oppositely situated with respect to the origin. Such oppositely situated points are (x,y,z) and $(-x,-y,-z)$ or (θ,ϕ) and $(\pi-\theta,\pi+\phi)$, as in Fig. 3.15. From the definition, YY^* involves even powers of $\cos \theta$ and $\sin \theta$; hence it is the same at these pairs of points for all Y 's. However, $Y_{l,m}$ is proportional to

$$e^{im\phi} (1 - \mu^2)^{|m|/2} \frac{d^{l+|m|} (\mu^2 - 1)^l}{d\mu^{l+|m|}}$$

where $\mu = \cos \theta$. $e^{im\phi}$ changes sign when $\phi + \pi$ is substituted for ϕ if m is odd and not if m is even. On performing the differentiation, one obtains a polynomial in μ of which the leading term is $\mu^{l-|m|}$ and subsequent terms are less by powers of μ^2 . Thus the polynomial changes sign or not with the substitution of $\pi - \theta$ for θ as $l - |m|$ is odd or even. The term $(1 - \mu^2)^{|m|/2} = \sin^{|m|} \theta$ does not change sign. Therefore $Y_{l,m}$ changes sign or not depending on whether $l - |m| + |m| = l$ is odd or not. Thus if l is even, $Y_{l,m}$ is an even function in the above sense and is said

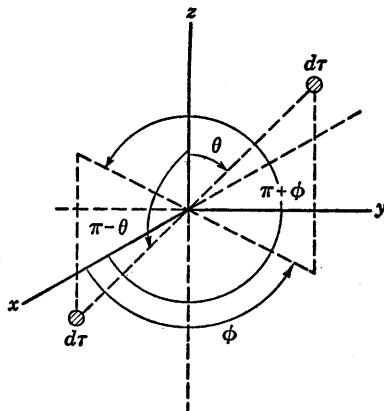


FIG. 3.15. Volume elements $d\tau$ symmetrically disposed about the origin.

to be of *even parity*. It is an odd function or said to be of *odd parity* if l is odd. Finally, the amplitude of $Y_{l,m}$ may be plotted as the length of a radius vector for all angles θ and ϕ , yielding lobed patterns if the terms $e^{im\phi}$ are replaced by $\sin m\phi$ and $\cos m\phi$. The number and placement of these lobes can be seen by inspecting the various individual functions, and they gain a particular significance in the consideration of molecular structure, as will be seen later. $Y_{0,0}$ is a sphere centered at the

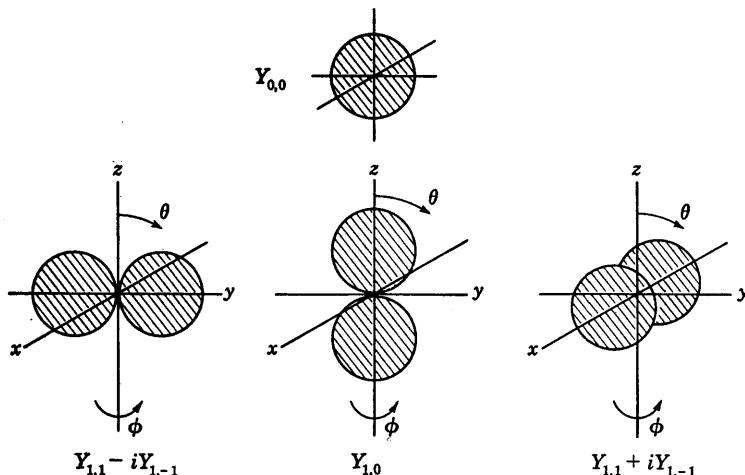


FIG. 3.16. Spherical or tesselar harmonic-function portrayals.

origin, and $Y_{1,0}$ and $Y_{1,1} \pm iY_{1,-1}$, also shown in Fig. 3.16, are spheres centered on the three axes and tangent to them at the origin.

The radial variation of u is described by Eq. (3.12). It is interesting to note that the parameter l from the angular equations enters as a term $l(l+1)/r^2$. If this term is multiplied by $\hbar^2/2m_r$, it has the dimensions of energy. If a force is associated with the negative gradient of this energy, it is $F = \hbar^2 l(l+1)/m_r r^3$. Now a centrifugal force is of the form $F = mr\omega^2$, which in terms of the angular momentum $L = mr^2\omega$ is

$$F = \frac{L^2}{mr^3}$$

Thus by comparison it appears suggestive to consider $\hbar^2 l(l+1)$ as a quantity corresponding to the square of the angular momentum associated with the motion. Though the uncertainty principle does not permit describing the system of the two particles in terms of their revolution as point masses about the common center of gravity, the concept of angular momentum retains a satisfactory definition and meaning as

$$L^2 = \hbar^2 l(l+1) \quad (3.13)$$

It is also consistent with the definition of angular momentum

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (3.14)$$

and the Schrödinger interpretation of momentum

$$\mathbf{p} = -i\hbar\nabla \quad (3.15)$$

This can be seen by forming the components of the angular-momentum vector:

$$\begin{aligned} L_x &= -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \\ L_y &= -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \\ L_z &= -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \end{aligned} \quad (3.16)$$

Expressing these in terms of the polar coordinates on a unit sphere, ϕ and θ , by using $\tan \phi = y/x$ and $\cos \theta = z(x^2 + y^2 + z^2)^{-\frac{1}{2}}$

$$\begin{aligned} L_x &= i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \\ L_y &= i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \\ L_z &= -i\hbar \frac{\partial}{\partial \phi} \end{aligned}$$

Then if L_x^2 means $-\hbar^2$ times the bracket repeated a second time, it is seen that

$$L^2 = L_x^2 + L_y^2 + L_z^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] = \hbar^2 l(l+1)$$

which, if considered as operating on $Y = \Theta\Phi$, is simply the equation resulting on the elimination of m from Eqs. (3.10) and (3.11). Thus in this sense Eqs. (3.16) represent the components of the angular momentum, the magnitude of which is

$$\hbar \sqrt{l(l+1)}$$

L_x and L_y are seen to be less simple in form in terms of the angles chosen than is L_z . In a sense this is fortuitous since it depended on the choice of angles with relation to the z axis, but there is also a real significance attached to the fact that it is possible to choose an axis for which the momentum component assumes this simple form. To illustrate this

consider the following expressions:

$$(xp_x - p_x x)u = -i\hbar x \frac{\partial u}{\partial x} + i\hbar \frac{\partial}{\partial x}(xu) = i\hbar u$$

$$(yp_x - p_xy)u = -i\hbar y \frac{\partial u}{\partial x} + i\hbar y \frac{\partial u}{\partial x} = 0$$

The variables x and p_x are complementary, and it is seen that the operator $(xp_x - p_x x)$, known as the *commutator*, does not vanish, that is, xp_x does not equal $p_x x$. On the other hand, yp_x does equal p_xy . It was seen in Sec. 2.5 that the quantities x and p_x cannot both be determined with arbitrary precision, whereas y and p_x can both be so determined. It can be shown in general that quantities that do not commute with one another (i.e., for which the order of multiplication is significant) are related in the complementary way, which is associated with their being independently measurable.

The components of angular momentum as given by Eqs. (3.16) are related in a somewhat similar way. Forming the indicated quantities it is seen that the following relations hold between them:

$$\begin{aligned} L_x L_y - L_y L_x &= i\hbar L_z \\ L_y L_z - L_z L_y &= i\hbar L_x \\ L_z L_x - L_x L_z &= i\hbar L_y \end{aligned} \quad (3.17)$$

Thus these components are related together in such a way that they can not each be independently determined with arbitrary precision. Equations (3.17) are an extension of Eq. (3.14) and prove to be more fundamentally descriptive of angular momentum than the classical definition. All atomic angular momenta are represented by Eqs. (3.17), whereas Eq. (3.14) represents the classical circulatory motions but not spin. Finally, it should be noted that the total angular momentum squared does commute with any one component so that it is possible to know the total angular momentum and a component along one axis at the same time without any limitation.

$$\begin{aligned} L_z L^2 - L^2 L_z &= L_z(L_x^2 + L_y^2 + L_z^2) - (L_x^2 + L_y^2 + L_z^2)L_z \\ &= (L_z L_x^2 - L_x^2 L_z) + (L_z L_y^2 - L_y^2 L_z) \\ &= i\hbar(L_y L_x + L_x L_y) - i\hbar(L_x L_y + L_y L_x) = 0 \end{aligned}$$

Because of the simplicity of the form of L_z in terms of angles, this is the component generally assumed to be determinable together with L^2 . To summarize

$$\begin{aligned} L^2 Y_{l,m}(\theta, \phi) &= \hbar^2 l(l+1) Y_{l,m}(\theta, \phi) \\ L_z Y_{l,m}(\theta, \phi) &= \hbar m Y_{l,m}(\theta, \phi) \end{aligned} \quad (3.18)$$

Equations (3.18) are seen to be closely similar in form to the quantum-mechanical energy equations such as (3.7) and (3.8). The physical interpretation is also the same as that adopted in the case of the energy equation; the state functions or eigenfunctions, which in this case are those known to describe the system with the energy E , also describe it with certain values of the angular momentum, and the differential operators on the left are identified with the particular values of the constants on the right for which the equations have the acceptable solutions $Y_{l,m}(\theta, \phi)$. Thus the operator \mathbf{l}^2 is equivalent to the total angular momentum squared, or $\hbar^2 l(l+1)$, and the operator \mathbf{l}_z , or $-\imath\hbar \partial/\partial\phi$, is equivalent to the component $m\hbar$ of angular momentum along a preferred axis where $|m| \leq l$. The close connection of this with the discussion of components of angular momenta and magnetic moments in earlier chapters is obvious.

Energy Quantization. The final step in the discussion of the hydrogen atom on the basis of Eq. (3.9) is the derivation of physically satisfactory solutions of Eq. (3.12). It is convenient to make the following substitutions:

$$R(r) = e^{-\rho/2} \rho^l L(\rho), \quad \rho = \frac{2Zr}{na_H}, \quad E_n = \frac{-Z^2\hbar^2}{2m_r n^2 a_H^2}$$

where $a_H = 4\pi\epsilon_0\hbar^2/m_r e^2$. These lead to

$$\rho \frac{d^2L}{d\rho^2} + [2(l+1) - \rho] \frac{dL}{d\rho} + (n-l-1)L = 0$$

A series solution $L(\rho) = \sum_k b_k \rho^k$ may be assumed for this equation (known as *Laguerre's equation*), and substitution shows that the following recurrence formula must hold for the b 's:

$$b_{k+1} = \frac{k-n+l+1}{(k+2l+2)(k+1)} b_k$$

For the series to terminate and assure the finiteness of R , n must be an integer, in which case the series terminates with b_k where $k = n - (l+1)$. Thus the conditions for a satisfactory solution are $n \geq (l+1)$ and $E_n = \frac{-m_r Z^2 e^4}{2(4\pi\epsilon_0)^2 \hbar^2 n^2}$, where n is an integer. The final component equation of (3.9) is thus solved, and the complete solution of the problem can be written.

To recapitulate, the solution of Eq. (3.6) is written as the product

$$\psi = \psi_c \psi_{n,l,m}$$

where $\psi_c = \psi_{c0} e^{-i(E_{ct} \pm p_c \cdot r)/\hbar}$ in which $p_c = \sqrt{2ME_c}$ represents the motion of the center of gravity, and

$$\psi_{n,l,m} = u_{n,l,m}(r, \theta, \phi) e^{-iE_n t/\hbar} \quad (3.19)$$

represents the motion of m_0 and M_p relative to the center of gravity. The component functions $\Phi_m(\phi)$, $\Theta_{m,l}(\theta)$, and $R_{n,l}(r)$ which make up u depend on certain integers m , l , and n and certain constant factors determined by the condition $\iiint u u^* r^2 \sin \theta dr d\theta d\phi = 1$. In terms of these factors so determined, the functions previously discussed are set down below:

$$\begin{aligned}\Phi_m(\phi) &= (2\pi)^{-\frac{1}{2}} e^{im\phi} \\ \Theta_{m,l}(\theta) &= \left[\frac{(2l+1)}{2} \frac{(l-|m|)!}{(l+|m|)!} \right]^{\frac{1}{2}} P_l^m(\cos \theta) \\ R_{n,l}(r) &= - \left\{ \left(\frac{2Z}{na_H} \right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3} \right\}^{\frac{1}{2}} \\ &\quad e^{-\frac{Zr}{na_H}} \left(\frac{2rZ}{na_H} \right)^l L_{n+l}^{2l+1} \left(\frac{2rZ}{na_H} \right)\end{aligned}$$

where

$$P_l^m(\cos \theta) = \frac{(\sin \theta)^{|m|}}{2l!} \left[\frac{\partial}{\partial(\cos \theta)} \right]^{l+|m|} (\cos^2 \theta - 1)^l$$

and

$$L_{n+l}^{2l+1} \left(\frac{2Zr}{na_H} \right) = \sum_{k=0}^{k=n-l-1} (-1)^{k+1} \frac{[(n+l)!]^2}{(n-l-k-1)!(2l+1+k)!k!} \left(\frac{2Zr}{na_H} \right)^k$$

The total energy can take values determined by

$$E_n = \frac{-m_r e^4 Z^2}{2(4\pi\epsilon_0)^2 \hbar^2 n^2} \quad \text{for } E < 0 \quad (3.20)$$

If $E > 0$, that is, the kinetic energy is greater than the energy of binding, the problem is a rather different one as the particles can get indefinitely far away from one another and no such condition as the above is necessary. Finally, the parameters n , l , and m are limited to the following values:

$n = 1, 2, 3, \dots, n$	<i>total quantum number</i>
$l = 0, 1, \dots, (n-1)$	<i>orbital quantum number</i>
$m = 0, \pm 1, \pm 2, \dots, \pm l$	<i>magnetic quantum number</i>

The most important condition on the solution of the radial portion of Eq. (3.9) is embodied in Eq. (3.20). This states that, if the kinetic energy of the electron is less than the potential energy of binding between the proton and electron so that they form a closed system, the total

internal energy of the system is limited to the discrete values determined by Eq. (3.20) with integral values for n . The dependence of E_n on the total quantum number correctly represents the Balmer type of term, and the hydrogenic spectrum series are properly described. Finally, the numerical value of the Rydberg constant for hydrogen ($Z = 1$) agrees with that predicted by Eq. (3.20). Since $E_n = h\nu_n = 2\pi\hbar c\bar{\nu}_n$,

$$\bar{\nu}_{n,m} = R_H \left(\frac{1}{m^2} - \frac{1}{n^2} \right)$$

where $R_H = m_r e^4 / 4\pi(4\pi\epsilon_0)^2 \hbar^3 c$.

Precisely the same theory with $Z = 2$ should apply to the system of one electron plus the helium nucleus, which is the ionized helium atom. Experimentally similar spectrum series are observed. The constant R_{He} of course differs from R_H both because of the difference in charge and the difference in reduced mass for the two cases. For a nucleus of very great mass and charge Z_e , the constant is

$$R_{\infty, Z} = \frac{m_0 Z^2 e^4}{4\pi(4\pi\epsilon_0)^2 \hbar^3 c} = \frac{\alpha^2}{2\lambda_e} Z^2$$

The wave numbers of the Balmer-series lines $n_i = 3 \rightarrow n_f = 2$ (H_α) and $n_i = 4 \rightarrow n_f = 2$ (H_β) for hydrogen and deuterium and the corresponding lines ($6 \rightarrow 4$) for once-ionized helium have been measured with great precision.¹ These measurements involve interferometric comparisons of the Balmer wavelengths with that of the standard red cadmium line. Cohen² has reevaluated the spectroscopic measurements of the H_α , H_β , D_α , and He wavelengths and has applied the relativistic, spin-orbit coupling, and Lamb-shift corrections to arrive at a value of $R_\infty = 10,973,732.6 \pm 1.4 \text{ m}^{-1}$.

The constant a_H , which has the dimensions of length, is known as the *first Bohr-orbit radius* for historical reasons. As a result of Rutherford's observations on the scattering of α particles, Bohr proposed a semi-classical model of a hydrogen atom in which the electron circulated in a planetary orbit about the nucleus. If the arbitrary condition that the angular momentum is limited to integral multiples of Planck's constant is then imposed, it is found that the permissible energy values are of the form of the Balmer terms with the correct value of the Rydberg constant and the radius of the smallest possible circular orbit is a_H . a_H is (m_0/m_r) times the first Bohr-orbit radius for infinite nuclear mass (a_0), and it is

¹ W. V. Houston, *Phys. Rev.*, **30**, 608 (1927); D. Y. Chu, *Phys. Rev.*, **55**, 175 (1939); C. F. Robinson, *Phys. Rev.*, **55**, 423 (1939); R. C. Williams, *Phys. Rev.*, **54**, 558, 568 (1938).

² E. R. Cohen, *Phys. Rev.*, **88**, 353 (1952).

interesting to note that

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_0e^2} = \frac{\lambda_c}{2\pi} \frac{1}{\alpha} = \frac{r_0}{\alpha^2}$$

where $\alpha = e^2/4\pi\epsilon_0\hbar c$ is the fine-structure constant, $\lambda_c = h/m_0c$ is the Compton wavelength, and $r_0 = e^2/4\pi\epsilon_0 m_0 c^2$ is the classical electron radius. Thus a_0 , which is about 0.53×10^{-10} m, is as much greater than the Compton wavelength divided by 2π (Sec. 2.4) as this quantity is greater than r_0 . The constant a_0 is a convenient measure of atomic distances, and the radial portion of the wave functions $R_{n,l}$ for a one-electron atom of infinite nuclear mass and charge Ze are most conveniently written in units of a_0/Z . These radial wave functions for a few values of n and l are plotted in Fig. 3.17 and listed as follows:

$$\begin{aligned} R_{1,0} &= 2 \left(\frac{Z}{a_0} \right)^{\frac{1}{2}} e^{-Zr/a_0} \\ R_{2,0} &= \frac{1}{\sqrt{8}} \left(\frac{Z}{a_0} \right)^{\frac{1}{2}} \left(2 - \frac{Zr}{a_0} \right) e^{-Zr/2a_0} \\ R_{2,1} &= \frac{1}{\sqrt{24}} \left(\frac{Z}{a_0} \right)^{\frac{1}{2}} \frac{Zr}{a_0} e^{-Zr/2a_0} \\ R_{3,0} &= \frac{2}{81\sqrt{3}} \left(\frac{Z}{a_0} \right)^{\frac{1}{2}} \left(27 - 18 \frac{Zr}{a_0} + \frac{2Z^2r^2}{a_0^2} \right) e^{-Zr/3a_0} \\ R_{3,1} &= \frac{4}{81\sqrt{6}} \left(\frac{Z}{a_0} \right)^{\frac{1}{2}} \left(6 - \frac{Zr}{a_0} \right) \frac{Zr}{a_0} e^{-Zr/3a_0} \\ R_{3,2} &= \frac{4}{81\sqrt{30}} \left(\frac{Z}{a_0} \right)^{\frac{1}{2}} \left(\frac{Zr}{a_0} \right)^2 e^{-Zr/3a_0} \end{aligned}$$

An atomic state for which $l = 0$ is known for historical reasons as an s state, one for which $l = 1$ is a p state, $l = 2$ is a d state, etc. The total quantum number is written as an integer before the letter, that is, $R_{1,0}$ corresponds to a $1s$ electron state, $R_{2,0}$ corresponds to a $2s$ electron state, $R_{2,1}$ corresponds to a $2p$ electron state. The squares of the radial portions of the wave functions are plotted in Fig. 3.17. If these are multiplied by r^2 and integrated from 0 to ∞ , the result is unity because the functions are normalized. It is interesting to note that R does not vanish in the neighborhood of the nucleus ($r = 0$) for s electron states. This is of importance in considering the interaction of a nuclear moment with an extranuclear electron, for such interaction will be particularly pronounced for s electron states.

A point of interest is that E_n , the energy of a characteristic level, depends on n alone of the three parameters n , l , and m . This is a peculiarity of the Coulomb field, and, if energies other than the electrostatic

energy between two monopoles are taken into account, E is found to depend on l as well as n . Thus in general an s electron will have a different energy for the same n than a p or d electron. As long as spherical symmetry of the Hamiltonian H is assumed, m will not appear in Eq. (3.12) and hence E will not depend on m . However, if the energy depends preferentially on one coordinate, H is no longer spherically symmetrical and E may then depend on all three parameters. Though the present discussion is limited to a single electron and a single nucleus

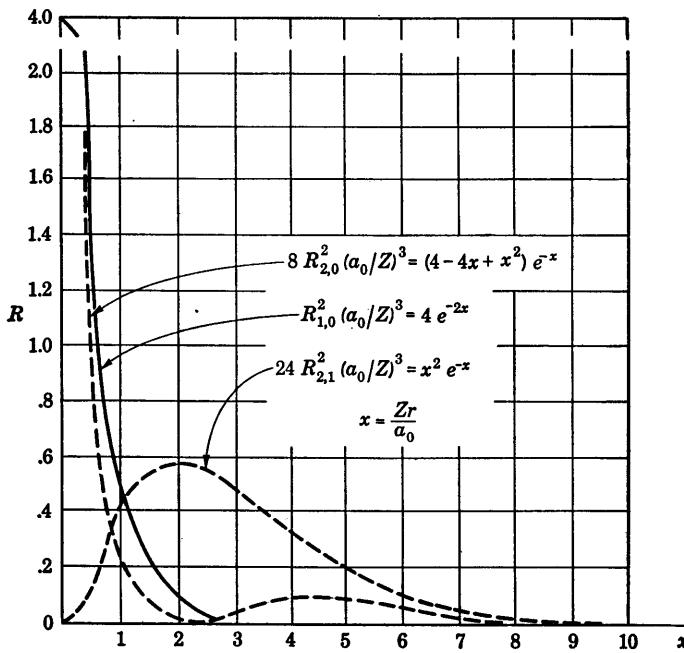


FIG. 3.17. Radial wave functions for a hydrogenic atom.

interacting only through their electric monopole charges, it includes the cases of singly ionized helium ($Z = 2$), doubly ionized lithium ($Z = 3$), etc. From the expressions for $R_{n,l}$ it is seen that the radial scale factor is inversely proportional to Z . Thus, as would be expected, the higher nuclear charge tends to draw the electron probability function farther in concentrating it near the nucleus resulting in a larger negative electrostatic energy for the system.

3.5. Quantum Theory of Radiation. Quantization of the Radiation Field. Since the primary means of excitation and deexcitation of atoms and molecules is the absorption and emission of electromagnetic radiation, the interaction of radiation with matter becomes of primary importance in atomic physics. The Bohr postulate that the quantum energy of

radiation represents the difference between energy levels of the atomic system allows the measurement of atomic energy levels. However, the complete investigation of atoms necessitates a more detailed theory of the radiation interaction. The Planck concept of the quantization of the electromagnetic radiation field offers a start toward the quantum-mechanical treatment of radiation. Nevertheless, development of the quantum theory of radiation was hampered by difficulties not encountered in atomic theory itself. The equations describing the interaction of the radiation field with even a single electron are too complicated for exact solution. A first approximation gives reasonable solutions.¹ Yet the second approximation results in integrals which diverge. By avoiding these divergencies the quantum electrodynamics gives predictions in good agreement with a large number of experiments. We shall develop in simple fashion the quantum-mechanical description of the interaction of atomic systems with electromagnetic radiation and apply it to several examples. First, we shall develop a suitable quantum-mechanical way of describing the radiation field.

The classical description of the electromagnetic radiation field may be conveniently expressed by the use of the vector potential \mathbf{A} . For the case of a region devoid of charges and assuming for simplicity that the scalar potential $\varphi = 0$, Maxwell's equations can be expressed in terms of \mathbf{A} as in Eq. (1.20),

$$\nabla^2 \mathbf{A} - \frac{\ddot{\mathbf{A}}}{c^2} = 0, \quad \nabla \cdot \mathbf{A} = 0$$

If the radiation field is enclosed in a volume, say L^3 , so that \mathbf{A} is periodic on the enclosing surface, then the solution of these equations can be gotten by the superposition of transverse plane waves of various wavelengths λ . By Fourier's theorem we may expand \mathbf{A} in a series of the form

$$\mathbf{A} = \sum_{\lambda} q_{\lambda}(t) \mathbf{A}_{\lambda}(\mathbf{r})$$

The variables t and \mathbf{r} are separable if $\ddot{q}_{\lambda} + \omega_{\lambda}^2 q_{\lambda} = 0$, where $\omega = 2\pi\nu$, so that

$$\nabla^2 \mathbf{A}_{\lambda} + \frac{\omega_{\lambda}^2}{c^2} \mathbf{A}_{\lambda} = 0$$

These are wave-type equations whose solutions can be written in the form

$$q_{\lambda} = s_{\lambda} e^{-i\omega_{\lambda} t}$$

¹ P. A. M. Dirac, "The Principles of Quantum Mechanics," 3d ed., chaps. 10, 12, Oxford University Press, London, 1947; W. Heitler, "The Quantum Theory of Radiation," 2d ed., chaps. 2, 3, Oxford University Press, London, 1944.

and

$$\mathbf{A}_\lambda = \mathbf{r}_{0\lambda} e^{i\mathbf{k}_\lambda \cdot \mathbf{r}}$$

where $\mathbf{r}_{0\lambda}$ is a unit vector indicating the polarization of the wave and \mathbf{K}_λ is the propagation vector $K_\lambda = \omega_\lambda/c$. $\mathbf{r}_{0\lambda}$ is perpendicular to \mathbf{K}_λ . The complete solution with two arbitrary constants can be written by including the complex conjugates which are also particular solutions of the wave equation.

$$\mathbf{A} = \sum_{\lambda} (q_\lambda \mathbf{A}_\lambda + q_\lambda^* \mathbf{A}_\lambda^*)$$

These waves \mathbf{A}_λ are orthogonal in the sense that $\int \mathbf{A}_\lambda \cdot \mathbf{A}_\mu^* d\tau = 0$, for $\lambda \neq \mu$, and $\int \mathbf{A}_\lambda \cdot \mathbf{A}_\lambda^* d\tau = 1$ as can be shown by Fourier's theorem. Because of the boundary conditions, the propagation vector \mathbf{K}_λ must be discrete:

$$K_{\lambda_x} = \frac{2\pi n_{\lambda_x}}{L}, \dots$$

The coefficients q_λ and q_λ^* can be combined to give real canonical variables Q_λ and their conjugates P_λ which are suitable for a Hamiltonian.

$$Q_\lambda = q_\lambda + q_\lambda^*$$

and

$$P_\lambda = -i\omega_\lambda(q_\lambda - q_\lambda^*) = \dot{Q}_\lambda$$

In order to determine the classical Hamiltonian for the radiation field in terms of the P_λ and Q_λ , we start with Eq. (1.23) for the energy of the field

$$U = \left(\frac{\epsilon_0}{2}\right) \int (E^2 + c^2 B^2) d\tau$$

and substitute $\mathbf{E} = -\dot{\mathbf{A}}$ and $\mathbf{B} = \text{curl } \mathbf{A}$ in terms of $\mathbf{A} = \sum_{\lambda} (q_\lambda \mathbf{A}_\lambda + q_\lambda^* \mathbf{A}_\lambda^*)$.

After vector expansions we obtain $U = \frac{\epsilon_0}{2} \sum_{\lambda} (P_\lambda^2 + \omega_\lambda^2 Q_\lambda^2)$. This is identical, formally, with the description of a set of harmonic oscillators. Although it is sometimes convenient to talk of these oscillators as if they were real and in equilibrium with the field, this is more of a formal mathematical than a physical analogy. Following the discussion of Sec. 1.8, the Hamiltonian for the oscillator is $H = \frac{1}{2}(\mathbf{p}_x^2/m + m\omega^2 x^2)$, which yields the wave equation

$$\left(\frac{\hbar^2}{m} \frac{d^2}{dx^2} + m\omega^2 x^2\right) \psi = W\psi$$

Later, in Sec. 5.4, it is seen that the eigenvalues of this equation are

$W_n = (n + \frac{1}{2})\hbar\omega$. In an assemblage of oscillators of different frequencies, the total energy will be the sum of the energy of each, $W = \sum_n W_n$.

The solutions of the harmonic-oscillator equation involve Hermite polynomials, and (as seen in Sec. 5.5) the position integrals

$$x_{nm} = \int_{-\infty}^{\infty} \psi_n x \psi_m^* dx$$

can be evaluated¹ to

$$x_{n,n+1} = \sqrt{\frac{n+1}{2\alpha}} \quad \text{where } \alpha = \frac{m\omega}{\hbar}$$

$$x_{n,n-1} = \sqrt{\frac{n}{2\alpha}}$$

$$x_{n,n'} = 0 \quad \text{for any other values of } n'$$

These results of quantizing the harmonic oscillator can be carried over to the quantization of the electromagnetic field. The constant $\frac{1}{2}$ can be dropped from W_n , since we will be concerned with differences in quantum energies. Hence the radiation field will have an energy $W = \sum_{\lambda} n_{\lambda} \hbar \omega_{\lambda}$.

The field can then be characterized by the quantum numbers n_{λ} describing the numbers of photons of frequency ω_{λ} . We can also apply the position-integral results to the radiation field problem by setting $m = 1$ and $s_{\lambda} = 1/\epsilon_0$ in the expressions above.

$$Q_{n,n+1} = \sqrt{\frac{\hbar(n+1)}{2\omega_{\lambda}\epsilon_0}} e^{-i\omega_{\lambda}t} = Q_{n+1,n}^*$$

$$Q_{n,n'} = 0$$

Or in terms of the complex amplitudes

$$q_{n,n+1} = \sqrt{\frac{\hbar(n+1)}{2\omega_{\lambda}\epsilon_0}} e^{-i\omega_{\lambda}t}$$

$$q_{n+1,n}^* = \sqrt{\frac{\hbar(n+1)}{2\omega_{\lambda}\epsilon_0}} e^{i\omega_{\lambda}t}$$

$$q_{n+1,n} = 0 = q_{n,n+1}^* = q_{n,n'}$$

Time-dependent Perturbation Theory. In order to calculate the results of the interaction between radiation and particles, we must assume that the interaction energy is small and find the effect of this perturbation on the probabilities of the system's being in the various stationary states

¹ L. Pauling and E. B. Wilson, "Introduction to Quantum Mechanics," sec. 11, McGraw-Hill Book Company, Inc., New York, 1935.

by seeking the approximate solutions to the first few orders in the interaction energy. This "time-dependent perturbation" theory is similar to the perturbation theory employed in Sec. 4.6 and later in Chap. 5.

The method of approximation which is here presented depends upon a property of solutions of equations such as (3.8), which is analogous to the property of the circular sine and cosine functions upon which the Fourier expansion of a function is based. This is the property of *orthonormality*. Consider the equations for two different allowed values of the energy E_n and E_m .

$$\frac{\hbar^2}{2m} \nabla^2 U_n - VU_n = -E_n U_n$$

$$\frac{\hbar^2}{2m} \nabla^2 U_m - VU_m = -E_m U_m$$

The complex conjugate of U_m , which is designated by an asterisk, is also a solution of that equation since both real and imaginary parts must separately be solutions and the sign of a solution can be changed without affecting the fact that it is a solution, i.e.,

$$\frac{\hbar^2}{2m} \nabla^2 U_m^* - VU_m^* = -E_m U_m^*$$

Multiplying the equation for U_n by U_m^* and that for U_m^* by U_n and subtracting, the term in V vanishes and

$$\frac{\hbar^2}{2m} (U_m^* \nabla^2 U_n - U_n \nabla^2 U_m^*) = (E_m - E_n) U_m^* U_n$$

But

$$\nabla \cdot (U_m^* \nabla U_n - U_n \nabla U_m^*) = (U_m^* \nabla^2 U_n - U_n \nabla^2 U_m^*)$$

so if the form on the left is used and the equation multiplied by a volume element and integrated, Stokes's theorem in the form $\int \nabla \cdot \mathbf{A} d\tau = \int \mathbf{A}_n \cdot d\mathbf{s}$ may be used to obtain

$$\frac{-\hbar^2}{2m} \int (U_m^* \nabla U_n - U_n \nabla U_m^*) \cdot d\mathbf{s} = (E_n - E_m) \int U_m^* U_n d\tau$$

where the surface s bounds the volume τ . Now the left-hand integral vanishes if the closed surface is so chosen that the U 's or their gradients vanish on this surface, or if the surface is at a very great distance from the system being described. For in the latter case the product $U \nabla U$ vanishes more rapidly than $1/r^2$, while the rate at which the area of the surface increases is only r^2 ; otherwise $\int U U^* d\tau$ would not be finite. Thus under either of the above conditions, one or the other of which can generally be brought about, the left side vanishes and therefore $\int U_n U_m^* d\tau = 0$ if

$m \neq n$. By suitable choice of the arbitrary constants contained in the U 's, one may set $\int U_n U_n^* d\tau = 1$, or in general

$$\int U_n U_m^* d\tau = \delta_{mn}$$

where $\delta_{mn} = 0$ if $m \neq n$, $\delta_{mn} = 1$ if $m = n$. This is the condition of orthonormality which is familiar for circular functions as well as for other harmonic functions. It enables one to expand certain functions in terms of the U 's just as functions are expanded in Fourier series in terms of sines and cosines.

We shall assume that the Hamiltonian of the system can be written as

$$H = H_0 + H'$$

where H_0 is the Hamiltonian of the unperturbed system whose eigenfunctions ψ_n are related to the eigenvalues E_n by $E_n \psi_n = H_0 \psi_n$. The perturbation H' is assumed to be small compared to H_0 . It will be assumed later that H' lasts a relatively short time. Then Schroedinger's equation is

$$-i\hbar \frac{\partial \Psi}{\partial t} = (H_0 + H')\Psi$$

We further assume that the wave-function solutions can be written

$$\Psi = \sum_n b_n(t) \psi_n e^{iE_n t/\hbar}$$

where the coefficients b_n are functions only of time and the ψ_n are eigenfunctions of the unperturbed system and functions only of the coordinates. The b_n 's are assumed to be normalized so that $|b_n(t)|^2$ is the probability of finding the system in the state with the eigenvalue E_n . Substituting this solution in Schroedinger's equation and subtracting out the unperturbed equation gives

$$\sum_n (-i)\hbar \psi_n \dot{b}_n(t) e^{iE_n t/\hbar} = \sum_n H' b_n(t) \psi_n e^{iE_n t/\hbar}$$

Now multiply by ψ_m^* and integrate over the coordinates. Since these functions are orthonormal,

$$\begin{aligned} \int \psi_n \psi_m^* d\tau &= 1 \text{ for } n = m \\ &= 0 \text{ for } n \neq m \end{aligned}$$

and

$$b_m(t) = \frac{i}{\hbar} \sum_n b_n(t) e^{i(E_n - E_m)t/\hbar} \int \psi_m^* H' \psi_n d\tau$$

If at time $t = 0$ the system is in the unperturbed state E_k , then $b_n(0) = 0$ except $b_k(0) = 1$. Let the perturbation H' last for a time t which is small

enough so that the b_n 's do not change appreciably but large compared to the "period" of the electron wave in the atom. Then at time t

$$-i\hbar \dot{b}_m(t) = (\int \psi_m^* H' \psi_k d\tau) e^{i(E_k - E_m)t/\hbar}$$

Let $H'_{mk} = \int \psi_m^* H' \psi_k d\tau$ and integrate with respect to t . If H' varies harmonically with time so that $H' = H'^0 e^{\pm i\omega t}$

$$-i\hbar \int_{b_m(0)=0}^{b_m(t)} d[b_m(t)] = H'^0_{mk} \int_{t'=0}^{t'=t} e^{i(E_k - E_m)t'/\hbar} dt' e^{\pm i\omega t'}$$

Let $\omega_{km} = (E_k - E_m)/\hbar$, so that

$$b_m(t) = H'^0_{mk} \frac{[e^{it(\omega_{km} \pm \omega)} - 1]}{\hbar(\omega_{km} \pm \omega)}$$

for $m \neq k$. Now the probability of finding the system in the state E_m is

$$|b_m(t)|^2 = |H'^0_{mk}|^2 \frac{2[1 - \cos(\omega_{km} \pm \omega)t]}{\hbar^2(\omega_{km} \pm \omega)^2}$$

This transition of the original system from state k to state m is associated with the emission or absorption of a photon of energy $\hbar\omega = E_k - E_m$. For small t this probability can be approximated to $(H'^0_{mk})^2 t^2 / \hbar^2$.

In some cases the matrix elements H'^0_{mk} for the transition from state k to state m vanish. In such a case there may exist some intermediate states g such that H'_{mg} and H'_{gk} do not vanish. Then the transition may proceed through these intermediate states, and the probability involves the sum over all possible states g .

In most radiation problems the energies for either the initial or the final states of the unperturbed system are positive, and the distribution of states is not quantized but continuous. So if we assume that the final state m is not a single state but consists of a large number of states of nearly the same energy E such that $\rho(E) dE$ is the number of states with energy between E and $E + dE$, then the transition probabilities must be multiplied by $\rho(E) dE$ and integrated over a small range ΔE . We also tacitly assume that H' depends only on E and not on the individual m 's, that is, that we can average over the angular momenta, etc., of the states m . Then if $\rho(E)$ is approximately constant over the range ΔE ,

$$\int \rho(E) \frac{1 - \cos(\omega_{km} \pm \omega)t}{\hbar^2(\omega_{km} \pm \omega)^2} dE = \rho(E_k) \pi t / \hbar$$

since $\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = \pi$ and the integral has its main contribution near $E_k = E_m + \hbar\omega_{km}$. The transition probability is then

$$|b_m(t)|^2 = |H'^0_{mk}|^2 \rho(E_k) \frac{2\pi t}{\hbar}$$

rather than $|H'_{mk}|^2 t^2/\hbar^2$ for transition to a single state. Schiff¹ indicates clearly how this transition probability to a set of states is associated with the uncertainty relation so that, when the perturbation is imposed, the energy of the system is quite uncertain, due to the disturbance of the perturbation. As time progresses, the transition, in a sense, is measuring the system energy by influencing the probability of finding the system with the energy E_k . For longer times, the precision with which E_k is known increases, thus narrowing the energy region into which the transition can proceed. This uncertainty in E_k then varies as \hbar/t by the uncertainty principle. Hence even though the transition probability to a single state E_k varies as t^2/\hbar^2 , the transition probability to a continuous set of levels near E_k will vary as $\frac{t^2}{\hbar^2} \frac{\hbar}{t}$ or t/\hbar , as indicated by the above integral. Condon and Morse² interpret the transition probability to a single state by considering a single photon or wave train. While this wave is acting on the system, the transition probability increases as t^2 . However, when this photon has passed, i.e., the wave train has passed, the probability stops increasing. Hence there is an average transition probability associated with each photon. The total probability will then equal this average value times the number of photons. This total probability will then increase linearly with the number of photons or linearly with time.

Now a transition probability per unit time can be written

$$T_{E_k, E} = \frac{1}{t} \int |b_m(t)|^2 \rho(E) dE = \frac{2\pi}{\hbar} \rho(E_k) |H'(E_k)|^2$$

where $H'(E_k) = H'_{mk}$ or $\sum_g \frac{H'_{mg} H'_{gk}}{\hbar(\omega_{kg} \pm \omega)}$ for intermediate states.

The transition probabilities are seen to be expressed in terms of the perturbation H'_{mk} which, of course, involves the perturbation Hamiltonian H' . The simplest interaction of the radiation field, i.e., with an electron, can be described in terms of the classical total energy. The Hamiltonian can be written as the sum of the energy of the radiation field plus the energy of the nonrelativistic charged particle (see Sec. 1.9).

$$H = \sum_{\lambda} n_{\lambda} \hbar \omega_{\lambda} + \frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2$$

¹ L. I. Schiff, "Quantum Mechanics," p. 191, McGraw-Hill Book Company, Inc., New York, 1949.

² E. U. Condon and P. M. Morse, "Quantum Mechanics," p. 242, McGraw-Hill Book Company, Inc., New York, 1929.

This can be expanded and approximated by neglecting the e^2 term and arranged as

$$H = H_0 + H' = \sum_{\lambda} n_{\lambda} \hbar \omega_{\lambda} + \frac{p^2}{2m} - \frac{e}{m} \mathbf{p} \cdot \mathbf{A}$$

where

$$H_0 = \sum_{\lambda} n_{\lambda} \hbar \omega_{\lambda} + \frac{p^2}{2m}$$

and

$$H' = \frac{-e}{m} \mathbf{p} \cdot \mathbf{A}$$

The H_0 term includes the energy of the radiation field and the kinetic energy of the particle. The H' term includes both the electron momentum and the vector potential and so describes the mutual interaction. The first-order interaction probabilities involve essentially $|H'|^2$ or, since we have seen that $H' \sim e$, the first-order processes involve e^2 . These are transitions involving only one photon emitted or absorbed. Processes in which two light quanta are involved must necessarily involve intermediate states for which the transition probabilities are proportional to $|H'|^4$ or e^4 . Thus radiation processes can be classified according to the number of photons involved or the power of $|H'|^2$ or e^2 .

First order, $T \sim e^2$, one photon involved: emission, absorption, one quantum annihilation of positrons, photoelectric effect.

Second order, $T \sim e^4$, two photons involved: dispersion, Raman effect, two quantum annihilation of positrons, Compton effect.

Third order, $T \sim e^6$, three photons (or one photon and a Coulomb deflection): bremsstrahlung, pair production.

Let us now consider in detail some simple first-order processes and calculate the transition probabilities.

3.6. Quantum Radiation Processes. Emission and Absorption of Photons by Atoms. We have seen that the Hamiltonian for an atomic system interacting with a radiation field can be written in terms of the unperturbed system with the atom in an initial state k and the radiation field described by the $n_1, n_2, \dots, n_{\lambda}, \dots$ quantum numbers plus the perturbation interaction. This perturbation will give a probability of transition to a new state m of the atom and a new set of quantum numbers of the field, $n'_1, n'_2, \dots, n'_{\lambda}, \dots$. This transition probability can be calculated as described in Sec. 3.5 when we know the matrix elements for the perturbation H' . Since we have seen that these matrix elements are zero unless a single n_{λ} changes to $n_{\lambda} + 1$, we can describe the emission of a photon by the quantum numbers

$$k, n_{\lambda} \rightarrow m, n_{\lambda} + 1$$

so that we wish to calculate the matrix $H'_{k,n_\lambda;m,n_\lambda+1}$. Let us assume that an atom in an excited atomic state E_k can make a transition to a lower atomic state $E_m + \hbar\omega$ and involves the continuum of states available for the photon.

From the preceding section

$$H'_{k,n_\lambda;m,n_\lambda+1} = \int \psi_k^* \left(\frac{-e}{m} \right) \mathbf{p} \cdot \mathbf{A} \psi_m d\tau$$

and, as has been shown,

$$\mathbf{A} = q_\lambda \mathbf{A}_\lambda + q_\lambda^* \mathbf{A}_\lambda^*$$

where $\mathbf{A}_\lambda = \mathbf{r}_{0\lambda} e^{i\mathbf{k}\lambda \cdot \mathbf{r}}$ and

$$q_{n_\lambda,n_\lambda+1} = \left[\frac{\hbar(n_\lambda + 1)}{2\omega_\lambda \epsilon_0} \right]^{\frac{1}{2}} e^{-i\omega_\lambda t}, \quad q_{n,n'} = 0$$

so

$$H' = \frac{-e}{m} \left[\frac{\hbar(n_\lambda + 1)}{2\omega_\lambda \epsilon_0} \right]^{\frac{1}{2}} \int \psi_k^* \mathbf{p} \cdot \mathbf{r}_{0\lambda} \psi_m e^{i\mathbf{k}\lambda \cdot \mathbf{r}} d\tau$$

The transition probability per unit time is then

$$T = \frac{2\pi}{\hbar} \rho(E_k) |H'|^2$$

The density of states near energy $E_k = E_m + \hbar\omega_{km} = E_m + E_{km}$ is determined by the discreteness of the allowed wavelengths of radiation in an enclosure and is shown in Sec. 6.5 to be

$$\rho(E) = \frac{8\pi E^2}{(2\pi c \hbar)^3}$$

The density of states of polarized radiation with angular frequency ω whose propagation vector is in the solid angle $d\Omega$ is

$$\rho(\omega) d\Omega = \frac{\omega^2 d\Omega}{(2\pi c)^3 \hbar}$$

Since T was derived by averaging over the final states, we must replace n_λ by \bar{n}_ω , the average number of photons of frequency ω with propagation vectors into $d\Omega$ and omit the subscript from ω . Then

$$T = \frac{e^2 \omega (\bar{n}_\omega + 1)}{m^2 2\pi \hbar c^3 4\pi \epsilon_0} \left| \int \psi_k^* \mathbf{p} \cdot \mathbf{r}_{0\lambda} \psi_m e^{i\mathbf{k}\lambda \cdot \mathbf{r}} d\tau \right|^2 d\Omega$$

If the wavelength $\lambda = 1/K_\lambda$ is large compared to the region where the wave functions of the particle have appreciable values, then $e^{i\mathbf{k}\lambda \cdot \mathbf{r}}$ is approximately 1. Let $\mathbf{p}/m = \mathbf{v}$ and let ϕ be the angle between $\mathbf{r}_{0\lambda}$ and \mathbf{v} ,

as in Fig. 3.18. Then

$$T = \frac{e^2 \omega (\bar{n}_\omega + 1) \cos^2 \phi}{2\pi \hbar c^3 4\pi \epsilon_0} \left| \int \psi_k^* v \psi_m d\tau \right|^2 d\Omega$$

If we choose the x axis along \mathbf{p} and θ , the angle between \mathbf{p} and \mathbf{K} , then $v_x = i\omega x$ and, averaging over the possible directions of polarization \mathbf{r}_ω for the chosen value of θ ,

$$T = \frac{e^2 \omega^3 \sin^2 \theta}{4\pi \hbar c^3 4\pi \epsilon_0} (\bar{n}_\omega + 1) \left| \int \psi_k^* x \psi_m d\tau \right|^2 d\Omega$$

Now $\int \psi_k^* e x \psi_m d\tau$ can be defined as the linear dipole moment P_z of a charge configuration. So

$$T = \frac{\omega^3 P_z^2 \sin^2 \theta}{4\pi \hbar c^3 4\pi \epsilon_0} (1 + \bar{n}_\omega) d\Omega$$

It will be noted that there are two parts to this transition probability. One is independent of the density of radiation n_ω and gives rise to "spontaneous emission." The other term is proportional to the radiation intensity at the frequency ω and produces induced emission. These two components of the emission probability were pointed out by Einstein on statistical grounds (see Sec. 6.4).

The angular distribution and the dependence on e , ω , and P are identical with the classical formula for dipole radiation. If the dipole moment is approximated by ea where a is the dimension of an atom, then the radiation probability is approximately $e^2 \omega^3 a^2 / \hbar c^3 (4\pi)^2 t_0$. For optical radiation this is of the order of 10^9 per second.

To determine the total intensity S spontaneously radiated per unit time, T can be multiplied by $\hbar\omega$ and integrated over a sphere.

$$S d\Omega = \frac{\omega^4 P^2 \sin^2 \theta d\Omega}{4\pi c^3 4\pi \epsilon_0}$$

$$S = \frac{\omega^4 P^2}{6\pi \epsilon_0 c^3}$$

The dipole radiation emitted when an atom changes from one energy state to another may also be described in terms of the classical theory of Sec. 1.7. The process may be thought of qualitatively as if the electric dipole moments of both states existed briefly together and the radiated frequency were the beat note between the characteristic frequencies $\omega_1 = E_1/\hbar$ and $\omega_2 = E_2/\hbar$ of the two states.

The effective dipole moment to be used is the average value of \bar{P} in the sense defined previously, but, instead of using the ψ functions for a single state, ψ is taken for one of the states and the conjugate ψ^* is taken

for the other of the two states between which the transition occurs. Thus

$$\bar{P}(t) = \int \psi_{n,l,m} \bar{\psi}_{n',l',m'}^* d\tau$$

is the equation for the radiating dipole moment on the above picture. The primed letters refer to one energy state and the unprimed letters to another. First of all, the time factor of the ψ functions yields immediately $e^{(-i/\hbar)(E_n - E_m)t} = e^{i\omega_{n,m}t}$, and the angular frequency is given properly by the difference in energy of the states divided by \hbar . Since the transition probabilities are determined by such effective moments, their evaluation is of interest. As an example consider the hydrogen atom as described in Sec. 3.4. The components of \bar{P} are given by

$$P_x = P_0 \sin \theta \cos \phi, \quad P_y = P_0 \sin \theta \sin \phi, \quad P_z = P_0 \cos \theta$$

where $P_0 = P_0(r)$. Writing the wave functions in terms of their separate functions R , Θ , and Φ ,

$$P_x [a_{n,l,n',l'}(r) b_{l,m,l',m'}(\theta) c_{m,m'}(\phi)]_x$$

and similarly for the y and z components, where

$$\begin{aligned} a_{n,l,n',l'} &= \int_0^\infty R_{n,l} P_0(r) R_{n',l'}^* r^2 dr \\ b_{l,m,l',m'} \left\{ \begin{array}{c} x \\ y \\ z \end{array} \right\} &= \int_0^\pi \Theta_{l,m} \left\{ \begin{array}{c} \sin \theta \\ \sin \theta \\ \cos \theta \end{array} \right\} \Theta_{l',m'}^* \sin \theta d\theta \\ c_{m,m'} \left\{ \begin{array}{c} x \\ y \\ z \end{array} \right\} &= \int_0^{2\pi} \Phi_m \left\{ \begin{array}{c} \cos \phi \\ \sin \phi \\ 1 \end{array} \right\} \Phi_{m'} d\phi \end{aligned}$$

These expressions are tedious to compute, but their general nature is of interest. The a 's do not vanish for any values of n or l , but the b 's and c 's vanish except for certain combinations; the conditions for their non-vanishing constitute the so-called selection rules for dipole radiation. Neglecting numerical factors,

$$c_{m,m'} \left\{ \begin{array}{c} x \\ y \\ z \end{array} \right\} = \int_0^{2\pi} e^{i(m-m')\phi} \left\{ \begin{array}{c} (e^{i\phi} + e^{-i\phi}) \\ (-i)(e^{i\phi} - e^{-i\phi}) \\ 1 \end{array} \right\} d\phi$$

so $c_{m,m'}(x)$ and $c_{m,m'}(y)$ vanish unless $m' = m \pm 1$; $c_{m,m'}(z)$ vanishes unless $m = m'$. These m selection rules relate to relative geometries and may be interpreted in terms of polarization. For transitions in which $m' \rightarrow m \pm 1$, $c_{m,m'}(x)$ and $c_{m,m'}(y)$ are $\pi/2$ out of phase, the angular-momentum component about the z axis changes; for conservation of angular momen-

tum this must represent angular momentum carried away by the radiation. In this case the radiation is circularly polarized if observed along z . The transitions $m \rightarrow m'$ involve no change of angular momentum about the z axis; hence the wave would appear plane-polarized with its electric vector along z . For a calculation of the b 's the definitions of Θ may be used to show that these coefficients vanish unless $l' = l \pm 1$. Therefore it is seen that the angular-momentum quantum number must change by exactly one unit. A p state may change to an s or d , etc., as was seen to be the case for the Grotrian diagrams of Sec. 3.1. From the discussion of the parity of the tesseral harmonics $Y_{l,m}$, the condition may be derived that in an electric dipole transition the total angular momentum must change from even to odd or the reverse, i.e., the parity of the system must change or the effective dipole moment for the transition is zero.

The treatment of *absorption* of electromagnetic radiation by a charge system is similar to that already described for emission. The intensity of a light beam in a given solid angle is

$$I(\omega) d\omega = \bar{n}_\omega \hbar \omega c \frac{\omega^2 d\omega}{(2\pi c)^3 \hbar} d\Omega$$

The transition probability is

$$T_a = (2\pi/\hbar)\rho(E)|H'|^2$$

where

$$H'_{m,n_\omega;k,n_\omega-1} = \frac{-e}{m} \sqrt{\frac{\hbar}{2\omega\epsilon_0}} \sqrt{\bar{n}_\omega} \int \psi_m \mathbf{p}_x e^{i\mathbf{k}\cdot\mathbf{r}} \psi_k d\tau$$

and

$$\rho(\omega) = \frac{\omega^2 d\Omega}{(2\pi c)^3 \hbar}$$

so

$$T_a = \frac{e^2}{m^2} \frac{\omega \bar{n}_\omega}{2\pi \hbar c^3 4\pi \epsilon_0} \left| \int \psi_m^* \mathbf{p} \cdot \mathbf{r}_0 e^{i\mathbf{k}\cdot\mathbf{r}} \psi_k d\tau \right|^2 d\Omega$$

Similarly, this can be reduced to

$$T_a = \frac{\omega^3 \bar{n}_\omega P_x^2 \sin^2 \theta d\Omega}{4\pi \hbar c^3 4\pi \epsilon_0}$$

The ratio of the emission probability to the absorption probability is

$$\frac{T_e}{T_a} = \frac{\bar{n}_\omega + 1}{\bar{n}_\omega}$$

The ratio will also be deduced from statistical arguments in Sec. 6.5. The total radiation absorbed per unit time is gotten by integrating T .

$$S = \frac{2\pi^2}{3} \frac{\omega I(\omega) P_x^2}{\hbar c 4\pi \epsilon_0}$$

The approximation that $e^{i\mathbf{K} \cdot \mathbf{r}} = 1$ in the integral of the wave functions for both emission and absorption is inexact if the wavelength of the radiation becomes comparable with the dimensions of the charge system. In such a case, the next approximation gives $e^{i\mathbf{K} \cdot \mathbf{r}} = 1 + i(\mathbf{K} \cdot \mathbf{r}) + \dots$. The perturbation matrix becomes $P_x + \int \psi_k^* iex \mathbf{K} \cdot \mathbf{x} \psi_m d\tau$. The second term is smaller than the first by the order of $(x/\lambda)^2$ unless the dipole moment vanishes. In this case the spontaneous transition probability includes only the second term

$$T_q = \frac{\omega^3 \sin^2 \theta d\Omega Q^2}{4\pi\hbar c^3 4\pi\epsilon_0}$$

where $Q = \int \psi_k^* iex (\mathbf{K}_0 \cdot \mathbf{x}) \psi_m d\tau$ is called the linear quadrupole moment. This is physically equivalent to considering the effects of retardation between different parts of the charge system. Here \mathbf{K}_0 is a unit vector directed to the point of observation.

Natural Line Breadth. In the time-dependent perturbation theory outlined in Sec. 3.5 we have assumed that the probability of the system's being in its original state is unchanged during the short time the perturbation is acting. If this were true the original state would be precisely known and the energy of the emitted photon precisely determined. If the perturbation has any appreciable effect, this will no longer be the case. The probability of the original state will decrease, the energy of the original state will be uncertain due to its finite mean lifetime τ , $\Delta E \tau \approx \hbar$, and the photon energy will be at least equally uncertain. This will contribute to a line breadth of the radiation observed. In order to calculate this effect let us reconsider the time-dependent perturbation equations with different auxiliary conditions.

$$\dot{b}_m(t) = \frac{i}{\hbar} \sum_n b_n(t) e^{i(E_n - E_m)t/\hbar} \int \psi_m^* H' \psi_n d\tau$$

At time $t = 0$ let $b_n(0) = 0$ except $b_k(0) = 1$. However, instead of letting $b_k(t) = 1$ we assume that it gradually decreases because of the perturbation as $b_k(t) = e^{-t/2\tau}$. Then

$$-i\hbar \dot{b}_m(t) = H_{mk}' e^{i(E_k - E_m - \hbar\omega)t/\hbar} e^{-t/2\tau}$$

and

$$b_m(t) = H_{mk}' \frac{e^{i(E_k - E_m - \hbar\omega)t/\hbar - t/2\tau} - 1}{E_k - E_m - \hbar\omega + i\hbar/2\tau}$$

Substituting this solution back into the equation for \dot{b}_k ,

$$\dot{b}_k = \frac{-e^{-t/2\tau}}{2\tau} = \frac{i}{\hbar} b_m H_{km}' e^{i(E_m - E_k + \hbar\omega)t/\hbar}$$

where the solution for τ is

$$\frac{1}{\tau} = \frac{-2i}{\hbar} |H'|^2 \frac{1 - e^{i(E_m - E_k + \hbar\omega)t/\hbar + t/2\tau}}{(E_k - E_m - \hbar\omega) + i\hbar/2\tau}$$

The right-hand side must be summed over the frequencies ω available, which is accomplished by integrating over $\rho \hbar d\omega d\Omega$, and, since $\int [(1 - e^{i(a-\omega)t+t/2\tau})/(a - \omega - i/2\tau)] d\omega = i\pi$,

$$\frac{1}{\tau} = \frac{2\pi}{\hbar} \rho \int |H'|^2 d\Omega$$

Now after a long time the probability that a given frequency has been emitted is

$$T(\omega) = |b_m|^2 = \frac{|H'|^2}{(E - \hbar\omega)^2 + \hbar^2/4\tau^2}$$

where $E = E_k - E_m$. The intensity of radiation is

$$\begin{aligned} I(\omega) d\omega &= \hbar\omega\rho\hbar d\omega \int |b_m|^2 d\Omega \\ &= \frac{1}{2\tau\pi} \frac{\hbar^3\omega d\omega}{[(E - \hbar\omega)^2 + \hbar^2/4\tau^2]} \end{aligned}$$

For $E - \hbar\omega = \hbar/2\tau$, the intensity is one-half its value at $E = \hbar\omega$ and so the full width at half maximum of the spectral line is \hbar/τ , the same as the uncertainty in energy of the photon due to the Heisenberg uncertainty relation. These derivations were made under the implicit assumption that the final state in the charge system was precise and had essentially infinite life. If, instead, the resultant state also decays, then its width must be combined with that of the initial state to determine the line width. Similarly, if the initial state can decay into other possible resultant states, their influence on the mean life of the initial state must be taken into account. In general, the natural line width of a transition from a to b will be

$$\frac{\hbar}{\tau_{ab}} = \sum_i \frac{\hbar}{\tau_{ai}} + \sum_j \frac{\hbar}{\tau_{bi}}$$

where the i and j states are those into which the states a and b can decay. An interesting test of this concept is afforded by the absorption in hydrogen of the H_α line ($n = 2$ to $n = 3$). The classical expression for line width involves only the intensity of emission and does not take into account the width of the resultant state. Unsöld¹ observed this absorption line in the solar atmosphere and measured the line shape. His

¹ A. Unsöld, *Z. Astrophys.*, **2**, 199 (1931).

results are in good agreement with the quantum-mechanical calculations and not with the classical predictions.

Photoelectric Effect. If the quantum energy of a photon is greater than the ionization energy I of an atom, the absorption of the photon can cause the ejection of the electron by the photoelectric effect. In this case the electron is raised into a state in the continuum and the kinetic energy of the electron is the difference between the photon energy and the ionization energy $\mathcal{E} = \hbar\omega - I$. This effect is still a first-order process and can be calculated by the methods of Sec. 3.5. The simplest case involves a photon of quantum energy much greater than the ionization energy of the electron concerned, but still nonrelativistic: $\hbar\omega \ll m_0c^2$, $\hbar\omega \gg I$.

In the initial state the electron is bound to the atom and its wave function can be assumed to be approximately hydrogenlike. As in Sec. 3.4, a 1s electron wave function is

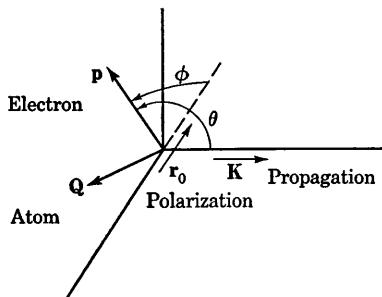


FIG. 3.18. Angles and vectors for the photoelectric effect.

$$\psi_a = \frac{1}{\pi^{\frac{1}{4}}} \left(\frac{Z}{a_0} \right)^{\frac{3}{2}} e^{-Zr/a_0}$$

where a_0 is the Bohr radius and the ionization energy I is

$$I = \frac{1}{2m_0} \left(\frac{Z\hbar}{a_0} \right)^2$$

The final state describes a free electron of momentum p whose wave function can be written

$$\psi_b = e^{ip \cdot r / \hbar}$$

This plane wave function is actually distorted by the long-range Coulomb field of the nucleus, so that a calculation using this plane wave is only an approximation. As indicated in the first part of this section, the level density for a free nonrelativistic particle, where $\lambda = h/p$ and $p^2/2m_0 = E$, is $\rho(E) dE d\Omega = pm_0 dE d\Omega/(2\pi\hbar)^3$. Let $\mathbf{K} = \mathbf{K}'\hbar c$ where \mathbf{K}' is the electron-wave propagation vector, and let $\mathbf{Q} = \mathbf{p} - \mathbf{K}/c$ be the momentum which the atom receives. Defining the angles as in Fig. 3.18, the projection of the electron momentum in the direction of polarization \mathbf{r}_0 is $p_e = p \sin \theta \cos \phi$. The perturbation matrix $H'_{a,1;b,0}^{''0}$ can be written as defined in Sec. 3.5 and integrated¹ and approximated for $K = \hbar\omega \ll m_0c^2$ to give

$$H' \cong \frac{-e}{m^3} \left(\frac{2\pi}{\epsilon_0} \right)^{\frac{1}{2}} \left(\frac{Z\hbar}{\omega a_0} \right)^{\frac{3}{2}} \frac{p \sin \theta \cos \phi}{(1 - \beta \cos \theta)^2}$$

¹ W. Heitler, *op. cit.*, p. 122.

where $\beta = v/c$. For $p \gg Z\hbar/a_0$, the differential cross section becomes

$$d\sigma_K = \frac{T}{c} = \frac{r_0^2 Z^5}{(137)^4} \left(\frac{m_0 c}{K}\right)^{\frac{1}{2}} \frac{4(2)^{\frac{1}{2}} \sin^2 \theta \cos^2 \phi d\Omega}{(1 - \beta \cos \theta)^4}$$

where r_0 is the classical electron radius.

The $(1 - \beta \cos \theta)$ term is negligible at low velocities of the photoelectron which then tends to be ejected parallel to the polarization direction, $\phi = 0^\circ$, $\theta = 90^\circ$. As $\beta = v/c$ becomes important, this angular distribution is tilted forward so that $\theta_{\max} \cong \pi/2 - 2\beta$. Neglecting this angular asymmetry and integrating over θ and ϕ gives the total photoelectric cross section for the two $1s$ electrons:

$$\sigma_K = \frac{r_0^2 Z^5}{(137)^4} \frac{32\pi(2)^{\frac{1}{2}}}{3} \left(\frac{m_0 c^2}{\hbar\omega}\right)^{\frac{1}{2}}$$

This cross section drops rapidly with photon energy but increases rapidly with Z . More refined calculations have been made by Hulme et al.¹ and others and summarized in Heitler² and by Davisson and Evans.³ Good agreement is shown between the more exact calculations and the experimental results of Allen⁴ and Gray⁵ on photoelectric absorption coefficients and those of Lutze⁶ and Williams et al.⁷ on angular distribution of the photoelectrons.

Dispersion and Raman Effect. Second-order processes in the interaction of radiation with matter involve two light quanta. Double photon emission and double photon absorption are negligible processes, but the scattering of light which is not negligible consists of the absorption of a quantum $\hbar\omega_0$ and the simultaneous emission of a quantum $\hbar\omega$. If the scattering system is left in its initial state, the scattering is coherent (or elastic); otherwise the scattering is inelastic or incoherent. If the photon energy is of the same order as the binding energy of the electron, the processes are called dispersion (elastic) and Raman effect (inelastic). For greater quantum energies the electron binding can be neglected and the process is that of Compton scattering.

Consider the scattering of light where the quantum energy is of the order of the electron binding energy, and identify the initial state as a and the final state b . Then $E_a - E_b = \hbar\omega - \hbar\omega_0$. In coherent scattering, $E_a = E_b$, while in Raman scattering, $E_a \neq E_b$.

¹ H. R. Hulme, J. McDouglass, R. A. Buckingham, and R. H. Fowler, *Proc. Roy. Soc. (London)*, **149A**, 131 (1935).

² W. Heitler, *op. cit.*, p. 124ff.

³ C. M. Davisson and R. D. Evans, *Rev. Mod. Phys.*, **24**, 79 (1952).

⁴ S. J. M. Allen, *Phys. Rev.*, **28**, 907 (1926).

⁵ L. H. Gray, *Proc. Cambridge Phil. Soc.*, **27**, 103 (1931).

⁶ E. Lutze, *Ann. Physik*, **9**, 853 (1931).

⁷ E. J. Williams, J. M. Nuttall, and H. S. Barlow, *Proc. Roy. Soc. (London)*, **121**, 611 (1928).

The perturbation Hamiltonian between photons and an electron was described in Sec. 3.5. Here we may not neglect the second-order term so

$$H' = -\frac{e}{m} \mathbf{p} \cdot \mathbf{A} + \frac{e^2}{2m} A^2 = H'_1 + H'_2$$

H'_1 can cause scattering through intermediate states, while H'_2 can cause scattering by direct transitions involving two quanta. The calculated¹ cross section is

$$d\sigma = r_0^2 \frac{\omega}{\omega_0} d\Omega \left(\frac{S}{mc^2} + \delta_{ab} \cos \theta \right)^2 \quad (\text{Raman scattering})$$

When $a = b$, coherent scattering occurs, $\omega = \omega_0$, and

$$d\sigma = r_0^2 d\Omega \left(\frac{S}{mc^2} + \cos \theta \right)^2$$

where

$$S = \sum_i \frac{P_{0ai} P_{ib}}{\hbar(\omega_{ai} + \omega_0)} + \frac{P_{ai} P_{0ib}}{\hbar(\omega_{ai} - \omega)}$$

and P_{0ai} is the projection of the matrix P_{ai} on \mathbf{K}_0 and P_{ai} is the projection on \mathbf{K} . θ is the angle between \mathbf{K}_0 and \mathbf{K} . These topics will be discussed at more length in later sections: coherent scattering in Sec. 5.1 and Raman scattering in Sec. 5.2.

Other Radiation Processes. When the quantum energy of the incident radiation equals a characteristic energy of the atom ($E_a - E_i$), the scattering becomes very large and is called *resonance fluorescence*. The dispersion formula described above actually goes to infinity when

$$\hbar\omega_0 = E_a - E_i$$

This is the result of neglecting in that formula the damping force or radiation resistance caused by the reaction of the emitted photon on the atom. A proper calculation of resonance fluorescence must take into account the finite lifetime of the excited states in a manner similar to that used previously for natural line breadth. This has been done by Weisskopf,² and the results are consistent with a coherent single quantum process as long as the atom is undisturbed during the process. If the quantum state of the atom is determined while scattering, as by a collision, etc., then the scattering is equivalent to an absorption and independent emission of a light quantum.

These scattering formulas were also limited by the assumption of strong binding of the electron, that is, $\hbar\omega_0 \approx E_i$. If the primary photon

¹ W. Heitler, *op. cit.*, pp. 129ff.

² V. Weisskopf, *Ann. Physik*, **9**, 23 (1931); *Z. Physik*, **85**, 451 (1933).

has an energy large compared to the ionization energy of the atom, the electron can be regarded as essentially free, and relativistic. In such a Compton scattering, momentum and energy are conserved between the incident photon, the scattered photon, and the electron, as mentioned previously. The intermediate states involved in this process now include the two spin directions of the electron and both its positive and negative energy states as described by the Dirac equation of Sec. 4.2. The summation is consequently more complicated and gives the Klein-Nishina¹ formula for the differential cross section of a photon of energy K_0 to be scattered by an electron into a solid angle $d\Omega$ at an angle θ when the incident radiation is unpolarized:

$$d\sigma_c = \frac{r_0^2 d\Omega}{4} \left(\frac{K}{K_0} \right)^2 \left(\frac{K_0}{K} + \frac{K}{K_0} - \sin^2 \theta \right)$$

where

$$K = \frac{K_0 m_0 c^2}{m_0 c^2 + K_0 (1 - \cos \theta)}, \quad r_0 = \frac{e^2}{m_0 c^2 4\pi \epsilon_0}, \quad K_0 = \hbar \omega_0$$

This distribution in angle of the Compton-scattered photons has been experimentally measured by Hofstadter and McIntyre² in good agreement with the theoretical formula. The total Compton cross section has been found by Parkinson³ to be in agreement with the Klein-Nishina predictions within 1 per cent. Cross and Ramsey⁴ have established the simultaneity of appearance of the Compton electron and photon to less than 1.5×10^{-8} sec and found the average angular correlation correct to 1°.

Processes involving three quanta are relatively less important since the probabilities are so small. However, the Coulomb deflection of an electron is equivalent to a second-order process in e^2 . Hence the emission of a photon due to the deceleration of an electron in the Coulomb field of a nucleus is a third-order radiation process. This continuous X radiation, called bremsstrahlung, is an important process, particularly in the production of cosmic-ray showers and continuous X rays. The cross section for the production of a photon of energy $\hbar\omega$ by an electron of energy E_0 in the field of a nucleus of charge Ze , for E_0 and $E > mc^2$, is

$$\sigma = \sigma_b \frac{8\pi}{\omega} \left[1 + \left(\frac{E}{E_0} \right)^2 - \frac{2}{3} \frac{E}{E_0} \right] \left(\ln \frac{2E_0 E}{mc^2 \hbar \omega} - \frac{1}{2} \right)$$

where

$$\sigma_b = \frac{Z^2 r_0^2}{137}, \quad E_0 = E + \hbar\omega$$

¹ O. Klein and Y. Nishina, *Z. Physik*, **52**, 853 (1929).

² R. Hofstadter and J. A. McIntyre, *Phys. Rev.*, **76**, 1269 (1949).

³ W. C. Parkinson, *Phys. Rev.*, **76**, 1348 (1949).

⁴ W. G. Cross and N. F. Ramsey, *Phys. Rev.*, **80**, 929 (1950).

In this calculation, the screening of the nuclear field by the atomic electrons has been neglected. This radiation loss by high-energy electrons in the cosmic rays gives rise to *showers* by a multiplicative process in which a very-high-energy electron produces a bremsstrahlung photon, the photon produces a pair (a positive and a negative electron), the pair particles again radiate, etc., and the number of charged particles builds up until ionization losses absorb the energy of the electrons and positrons. The experimental confirmation of this phenomenon gives reliable evidence that the theory is correct even at very high energies.

The creation of a pair (a positive and a negative electron) by gamma rays in the field of a nucleus is likewise a third-order process since it involves one photon and the interaction of an electron with the Coulomb field of a nucleus. It is very similar to bremsstrahlung but in reverse and with a negative energy state for the final state of the electron. Consequently, the pair-production cross section involves the σ_b mentioned in the bremsstrahlung formula but has an energy threshold, 1.02 Mev, to create the masses of the positron and electron. Pair production may be thought of as the absorption of a photon in the field of a nucleus with the excitation of an electron from a negative energy state to a positive energy state. The "hole" left behind acts like an electron of opposite sign, i.e., a positron (see Sec. 4.3).

A positron can annihilate in conjunction with an electron by a process described as the electron falling into the "hole" in the negative energy states. Normally, two quanta are emitted in order to conserve momentum. The annihilation of positrons in matter is complicated by the formation of positronium, an atomlike combination of a positron and an electron. Positronium has a measurable lifetime, and its quantum states can be observed.¹

PROBLEMS

- 3.1.** Consider the scattering of an α particle by a Thomson-type atom in which the positive atomic charge is uniformly distributed over a spherical volume whose radius R is so large that the number of α particles passing through the atom cannot be neglected. Show that in this case the small deflection experienced by an α particle passing a distance b from the center of the atom is

$$\frac{Zee'}{2\pi\epsilon_0 p_s v_z} \frac{1}{b} \left[1 - \left(1 - \frac{b^2}{R^2} \right)^{\frac{1}{2}} \right]$$

for b less than R . Show that the mean-square deflection in the plane of the trajectory per collision for particles being scattered by such an atom is

$$\left(\frac{Zee'}{2\pi\epsilon_0 p_s v_z b_{\max}} \right)^2 2 \left[\ln \frac{b_{\max}}{R} + (\frac{3}{4} - \ln 4) \right]$$

where the letters have the same significance as in the text.

¹ M. Deutsch, Annihilation of Positrons, vol. 3, of O. R. Frisch, "Progress in Nuclear Physics," Academic Press, Inc., New York, 1953.

3.2. The discussion of scattering by a Thomson-type atom in which the ultimate deflection is produced by a large number of small individual deflections can be employed to yield an approximate theory of the scattering of fast electrons by matter. The quantity b_{\max} may be taken as of the order of magnitude of atomic dimensions, i.e., $5 \times 10^{-11} Z^{-\frac{1}{3}}$ (Sec. 6.5), where Z is the atomic number. The quantity b_{\min} may be taken as determined by the assumption that no individual deflection is large, that is

$$b_{\min} \cong \left(\frac{1}{\alpha}\right) \frac{Zee'}{2\pi\epsilon_0 p_z v_z}$$

where α is the fine-structure constant. Alternatively it might be determined by saying that the distance of closest approach to a scattering center is not less than the effective wavelength of the electron. Show that for electrons in the million-ev energy range these conditions are essentially equivalent and b_{\min} can be taken as the Compton wavelength λ_c . Show that the mean-square deflection of an electron in the million-ev energy range on passing through a thickness t meters of material of density d in kg m^{-3} , atomic number Z , and atomic weight A is of the order of

$$\overline{\phi^2} \cong 10^{11} \left(\frac{Z}{V}\right)^2 \frac{d}{A} t$$

where V is the energy of the electron in electron volts. Determine the thickness of air and of lead for which $\overline{\phi^2}$ would be of the order of unity for million-ev electrons and hence the emerging electrons would occur in practically the entire forward hemisphere.

3.3. An elastic collision between a particle of mass M_1 traveling with a velocity v_1 and a particle of mass M_2 initially at rest is analyzed in two coordinate systems: (1) the laboratory system, and (2) that for which the center of gravity is at rest. Show that the second system may be obtained from the first by applying the constant-velocity vector $-M_1 v_1 / (M_1 + M_2)$ to both particles and that the angular deflection of the particles M_1 in the second system (θ') is related to its deflection in the laboratory system (θ) by

$$\cot \theta = \cot \theta' + \frac{M_1}{M_2} \csc \theta'$$

3.4. Consider many collisions of the type in the preceding problem and note that the number of scattered particles must be the same when described in terms of either coordinate system, that is, $\int \sigma(\theta, \phi) d\omega = \int \sigma(\theta', \phi') d\omega'$, where θ and ϕ are the polar and azimuthal angles about the oncoming particles, $d\omega$ is the solid angle $\sin \theta d\theta d\phi$, σ is the scattering cross section, and the primes refer to the coordinate system for which the center of gravity is at rest ($\phi = \phi'$). In consequence of this and the result of the preceding problem, show that the equation for the elastic-scattering cross section in the laboratory system in terms of the cross section in the primed system is given by

$$\sigma(\theta) = \sigma'(\theta') \left[1 + 2 \frac{M_1}{M_2} \cos \theta' + \left(\frac{M_1}{M_2} \right)^2 \right]^{\frac{1}{2}} \left(1 + \frac{M_1}{M_2} \cos \theta' \right)^{-1}$$

3.5. Using the result of Prob. 3.3 show that if M_1 is less than M_2 the energy of each particle in the laboratory system after the scattering process is a single-valued function of the angle which its trajectory makes with v_1 . Show that if M_1 is greater than M_2 the above statement is true for M_2 but that the M_1 particles do not appear for angles greater than θ_m where $\sin \theta_m = M_2/M_1$ and that for angles less than θ_m there are two characteristic energies for these particles. Show also that the energies relative to the

initial energy of M_1 are given by

$$E_{M_1} = \frac{M_1 M_2}{(M_1 + M_2)^2} \left(\frac{M_1}{M_2} + 2 \cos \theta' + \frac{M_2}{M_1} \right)$$

$$E_{M_2} = \frac{M_1 M_2}{(M_1 + M_2)^2} 2(1 - \cos \theta')$$

and the distance of closest approach by

$$r = r_{\min} \frac{1 + \sin(\theta'/2)}{2 \sin(\theta'/2)}$$

θ' is the scattering angle of M_1 in the system in which the center of gravity is at rest, and r_{\min} is the distance of closest approach for a head-on collision (inverse-square repulsion).

3.6. A collimated beam of polonium α particles, for which the energy is 6 Mev, is incident upon a polyethylene (C_2H_4) film of such a thickness that 1 m^2 of it weighs 1 g. The scattered and recoil particles can enter the window of a crystal counter which subtends an angle of 10^{-3} steradian at the scattering area. The geometry is such that the angle between the incident beam and the trajectory of a particle entering the counter can be set at the four values 10° , 30° , 60° , and 120° . Assume that 10^6 particles impinge on the film per second and that the counter records the energy as well as the number of particles of all types entering it. Determine the number and energy of the different types of particles observed by the counter at each of the four angles. Also determine the least distances of approach for the α particles and the two nuclei. The Coulomb law is assumed to hold in the above calculations.

3.7. In Sec. 3.2 the small angular deflection due to the electrostatic forces between a charged particle and a charged scattering center when they pass at a distance of closest approach b was derived. Assuming that the scattered particle possesses a magnetic moment in addition to its charge and is of such a character that its magnetic energy in a magnetic induction B can take only the two values $\pm e\hbar B/2m_0$, show that the resulting small angular deflection due to both electrostatic and electromagnetic forces is equal to $(1 \pm \lambda v_z / 4\pi bc)$ times that due to the electrostatic forces alone. Here λ_c is the Compton wavelength, v_z is the particle velocity, b is the distance of closest approach, and c is the velocity of light.

3.8. Assuming the magnetic induction to be uniform over the region in which the perturbing electromagnetic field is produced in an atomic beam measurement, show that the precision $(\delta\omega/\omega)$ with which γ may be determined is of the order of ρ/l , where ρ is the radius of curvature of motion of the atoms if they were ionized and moving in the magnetic induction B , and l is the length of the beam over which the perturbing field is applied. Assuming $l = 10^{-1} \text{ m}$, $B = 1 \text{ weber m}^{-2}$, the mean velocity of atoms in the beam is 10^3 m sec^{-1} and their mass is 10 atomic-weight units; calculate $\delta\omega/\omega$.

3.9. Consider the passage of light through a 60° prism at *minimum deviation*, which is the condition in which the light rays in the prism are parallel to the base. If θ is the total angle through which the wave normal is turned on passing through the prism, n is the index of refraction of the prism material, and a is the aperture or projected length of a prism edge t at right angles to the wave normal, show that $t dn/d\lambda = a d\theta/d\lambda$. The first diffraction minimum for the aperture a occurs at the angle $\Delta\theta = \lambda/a$; thus the resolving power, which is λ divided by the smallest detectable change in λ , is $\lambda/\delta\lambda = t dn/d\lambda = a d\theta/d\lambda$. Using the data given in the text for glass, show that if a grating has 600 lines per centimeter, its resolving power in the first order is the same as that for a 60° glass prism with a base length equal to the length of the ruled area of the grating.

3.10. The frequency difference between the two fine-structure components of the hydrogen spectrum line H_α at 6562 \AA is about 1000 Mc. How low would the tem-

perature of a gas-discharge source have to be for direct observation to resolve the component of H_α for (a) hydrogen, (b) deuterium? If the collision radius is 10^{-10} m for interrupting the radiation process, how low a pressure would be necessary in the discharge emitting the spectrum for resolution in the two cases of hydrogen and deuterium? If the discharge is not observed directly but a narrow beam of atoms emerging from the discharge is viewed normally by the spectroscope, show that the usable temperature is increased by a factor of the order of the square of the cotangent of the semiconical angle of the beam. Using the optical values in the text, show that a glass prism capable of resolving this fine structure would have to be about 7.65 m on a side.

3.11. A fundamental resonant absorption frequency of the molecule ICl^{35} is found to occur at 6980 Mc. Radio-frequency power from a well-stabilized oscillator in this frequency range is supplied to a cavity in connection with which there is a square-law detector, the response of which is proportional to the power level in the cavity. The Q value of the cavity when evacuated and tuned to the ICl^{35} frequency is found to be 30,000, and when ICl^{35} is admitted to the cavity at a pressure of 1 cm of mercury the power level is found to drop by 4.1 per cent. Show that the absorption or attenuation coefficient of ICl^{35} for radiation under these circumstances is 10^{-4} m^{-1} or that the effective absorption cross section is $2.8 \times 10^{-28} \text{ m}^2$ per molecule.

3.12. Assume that the experiment of Prob. 3.11 was performed at room temperature ($T = 300^\circ \text{ abs}$) and that the ratio of the populations of the upper and lower states concerned in the transition is given by the Boltzmann factor $e^{-h\nu/kT}$. Using the results of Prob. 3.11 and assuming that 1 per cent of the molecules present are known to be in the states concerned, show that the cross sections of the molecule for absorption and induced emission (taking these to be the same) are given at resonance by $2.5 \times 10^{-23} \text{ m}^2$.

3.13. As it has been seen that an amount of angular momentum equal to \hbar leaves an atom in a dipole-radiation process, show that this is consistent with the statement that if the photon has no intrinsic angular momentum or spin it must leave the atom normal to a radius vector from the center of the atom at a distance equal in magnitude to the wavelength of the radiation divided by 2π . Compare the radiating cross section thus implied with the resonant scattering cross section of an oscillator as given in Prob. 1.16.

3.14. Show that the radial portions of the hydrogenlike wave functions for a heavy atom of atomic number Z which do not vanish at $r = 0$ are those for which $l = 0$, and that the effective electron-charge density at the origin due to such functions is

$$(euu^*)_{r=0} = \frac{e}{\pi} \left(\frac{Z}{na_0} \right)^3$$

From Poisson's equation ($\nabla^2\varphi = -\rho/\epsilon_0$), the condition of spherical symmetry, and the definition of the nuclear quadrupole moment, show that the energy of interaction between an s electron and a nuclear quadrupole moment Q is

$$U_Q = \frac{c\hbar\alpha Z}{3} \left(\frac{Z}{na_0} \right)^3 Q$$

where α is the fine-structure constant.

3.15. The wave number of one of the close components constituting the red line of the hydrogen spectrum H_α ($n = 3$ to $n = 2$) has been measured and found to be $15,233.094 \text{ cm}^{-1}$ with a precision of about 0.2 part per million. The wave-number difference between this component and the analogous component of the deuterium spectrum D_α is found to be $\Delta B = 4.1461 \text{ cm}^{-1}$ with a precision of about 100 parts per million. Using the values of the masses of the hydrogen and deuterium ions and the

above data show that the electron mass is found to be 5.4899×10^{-4} in atomic-weight units with a precision of about 1 part in 10^4 .

3.16. In this chapter the contribution to the atomic energy levels arising from the interactions of the intrinsic magnetic moment of the electron with the moment arising from its orbital motion has been neglected. Assume that as a first approximation the magnetic moment arising from the angular momentum of the electron's motion can be represented by a moment $l\epsilon\hbar/2m_0$ at the nucleus and that the two components of the spin moment at a distance na_0/Z normal to l are $\pm\epsilon\hbar/2m_0$. Considering an atom consisting of one electron and a nucleus of charge Z and using the expression for the energy of two magnetic dipoles from Chap. 1, show that the difference between the fine-structure components of a level of quantum number n and orbital quantum number l is $\frac{1}{2}m_0c^2l\alpha^4(Z/n)^3$ and that the ratio of this energy to that of electrostatic origin is $l\alpha^2Z^2/n^2$. α is the fine-structure constant. Compare this order of magnitude with the observed fine structure of the following emission lines: (a) the hydrogen line H_α ($n = 3$ to $n = 2$) for which $\Delta\bar{\nu} = 0.3 \text{ cm}^{-1}$, and (b) the helium-ion line at

$$\lambda = 4686 \text{ A} \quad (n = 4 \text{ to } n = 3)$$

for which $\Delta\bar{\nu} = 0.45 \text{ cm}^{-1}$. The orbital quantum number l may be taken as unity in both cases.

3.17. From the argument that the parity of $Y_{l,m}$ depends only on whether l is even or odd, show that the angular function $b_{l,m,l',m'}(\theta)c_{m,m'}(\phi)$ of the polarization vanishes unless $Y_{l,m}$ and $Y_{l',m'}$ are of different parity. Thus there is no electric dipole radiation for a transition between states of like parity regardless of the dependence of the wave functions on r . Using the discussion of Sec. 2.7 show that magnetic dipole and electric quadrupole radiation is forbidden between states of unlike parity.

3.18. The existence of a particle with the same mass and spin as an electron and with a charge equal in magnitude but opposite in sign was established by experiments of Anderson. This particle, which is brought into existence as a result of high-energy collision processes is called a *positron*. It has but a brief existence because under the proper conditions it can merge with an electron and disappear, and the mass energy of the pair is emitted as radiation. During its existence, however, the positron and electron can form an ephemeral hydrogenlike system called *positronium*, in which the positron takes the part of the proton in hydrogen. Show that the energy levels of positronium are analogous to those of hydrogen, that their magnitudes are less by the factor $\frac{1}{2}(1 + m_0/M_e)$, and that the linear parameter measuring the size of the system is greater by the reciprocal of this factor than for hydrogen.

3.19. A convenient system of units for atomic calculations is one in which the elementary magnitudes are taken as unity. In one such system of atomic units proposed by Hartree the unit of mass is the electron rest mass m_0 , the unit of length is the first Bohr radius a_0 , and the unit of angular momentum is \hbar . Show that in this system the unit of velocity is αc , where α is the dimensionless fine-structure constant and c is the velocity of light, and that the unit of energy is twice the energy of the ground state of a hydrogenic atom with a single nuclear charge and an infinite mass.

3.20. Calculate the contributions of Doppler shift, collisions, and natural lifetime to the line width of the absorption line H_α in the hydrogen in the atmosphere of the sun. Assume the temperature to be about 6000°K , the pressure of the order of 10^{-3} atm. The transition probabilities from $n = 3$ to $n = 2$ are about 9×10^7 per second, from $n = 3$ to $n = 1$ about 17×10^7 , and from $n = 2$ to $n = 1$ about 63×10^7 .

3.21. Calculate and plot the energy and angular distributions of the photoelectrons ejected from a thin gold foil by the 0.279-Mev gamma rays of Hg^{203} . Use the non-relativistic approximation. Assuming that a source of 1 millicurie of Hg^{203} is available, what are the actual numbers of photoelectrons that might be observed?

3.22. Statistical mechanics (Chap. 6) gives the ratio of excited atoms due to thermal excitation at a temperature T to those in the ground state to be

$$\frac{N_e}{N_g} = e^{(W_e - W_g)/kT} = e^{-\hbar\omega_{eg}/kT}$$

Assuming the Einstein coefficient of induced emission to be equal to the Einstein coefficient of absorption, show that induced emission is equal to spontaneous emission at $T = \hbar\omega/k \ln 2$.

3.23. Show that the selection rules forbid a hydrogen atom in a rectangular box to radiate any of its translational kinetic energy.

3.24. Show that $\sum_{m=-l}^{-l} Y_{l,m} Y_{l,m}^*$ is independent of angle for $l = 2$.

3.25. Carry through the calculation of the photoelectric effect outlined in Sec. 3.6.

GENERAL REFERENCES

- Bacher, R., and S. Goudsmit: "Atomic Energy States," McGraw-Hill Book Company, Inc., New York, 1932.
- Compton, A. H., and S. K. Allison: "X-Rays in Theory and Experiment," D. Van Nostrand Company, Inc., New York, 1935.
- Condon, E. U., and P. M. Morse: "Quantum Mechanics," chaps. 1, 2, McGraw-Hill Book Company, Inc., New York, 1929.
- Condon, E. U., and G. H. Shortley: "Theory of Atomic Spectra," Cambridge University Press, London, 1935.
- Dirac, P. A. M.: "Quantum Mechanics," chaps. 10, 12, Oxford University Press, London, 1944.
- Fraser, R.: "Molecular Beams," Methuen & Co., Ltd., London, 1937.
- Gaydon, A. G.: "Dissociation Energies," Chapman & Hall, Ltd., London, 1947.
- Harrison, G. R., R. C. Lord, and J. R. Loofbourrow: "Practical Spectroscopy," Prentice-Hall, Inc., New York, 1948.
- Heitler, W.: "The Quantum Theory of Radiation," chap. 3, Oxford University Press, London, 1944.
- Herzberg, G.: "Atomic Spectra and Atomic Structure," chaps. 1, 2, Dover Publications, New York, 1944.
- Houston, W. V.: "Principles of Quantum Mechanics," McGraw-Hill Book Company, Inc., New York, 1951.
- March, A.: "Quantum Mechanics of Particles and Wave Fields," chap. 8, John Wiley & Sons, Inc., New York, 1951.
- Mott, N. F., and I. N. Sneddon: "Wave Mechanics," Oxford University Press, London, 1948.
- Pauling, L., and E. B. Wilson: "Introduction to Quantum Mechanics," chaps. 3, 4, 5, McGraw-Hill Book Company, Inc., New York, 1935.
- Rutherford, E., J. Chadwick, and C. D. Ellis: "Radiations from Radioactive Substances," Cambridge University Press, London, 1930.
- Sawyer, R. A.: "Experimental Spectroscopy," Prentice-Hall, Inc., New York, 1944.
- Schiff, L. I.: "Quantum Mechanics," chaps. 1-4, 8, 10, 13, McGraw-Hill Book Company, Inc., New York, 1949.
- White, H. E.: "Introduction to Atomic Spectra," chaps. 1-4, 21, McGraw-Hill Book Company, Inc., New York, 1934.

CHAPTER 4

ELECTRON SPIN AND POLYELECTRONIC ATOMS

4.1. Concept of Spin and Simple Magnetic Effects. In Sec. 3.4 it was shown that the classical concepts of energy, momentum, and angular momentum are adequate to describe the principle features of the hydrogen atomic spectrum in concordance with the wave nature of the constituent elementary particles provided that momentum is interpreted as the operator $\mathbf{p} = -i\hbar\nabla$ and energy as the operator $E = i\hbar \partial/\partial t$. However, if sufficient spectroscopic resolution is employed, the apparently simple spectrum lines may be distinguished as a number of lines close together, called a spectrum multiplet. The classical concepts of charged particles in rotational motion are inadequate to account for the types of phenomena that are observed. This failure might be thought to be due to the neglect of the interaction of the magnetic moment of a possible classical rotational motion of the electron about its own center of gravity with the magnetic moment of the electron's motion of revolution about the nucleus. However, the result of a calculation including this interaction does not agree with the observed first-order multiplicity. The interaction of electronic moments with the moment of the nucleus has been seen to be several orders of magnitude smaller than the observed splitting, so it cannot be invoked to provide an explanation.

It is readily seen from the discussion of Sec. 3.4 that the classical magnetic interaction energy between an intrinsic electronic rotational motion and the rest of the atomic system would lead to an odd number of energy fine-structure components and three lines in a spectrum multiplet. The components of angular momentum of a rotational or spinning motion along some chosen axis such as that of the orbital motion are, by Eq. (3.18), $m\hbar$, where m can be zero or a range of positive and negative integral values. Assuming the energy of interaction to be of the magnetic type and the magnetic moment to be in the classical ratio to the angular momentum, the energy levels would consist of three components ($m = -1, 0, 1$) or five components ($m = -2, -1, 0, 1, 2$), etc., for successively higher values of the angular momentum. By the selection rule, $\Delta m = 0, \pm 1$, the energy changes during an emission or absorption process in the presence of a magnetic induction \mathbf{B} may be either unchanged by

the magnetic induction or greater or less by the quantity $(e\hbar/2m_0)B$. Thus a single spectrum line would be changed by the magnetic induction to a triplet. Very high spectroscopic resolution or microwave techniques show, in contradiction to this, that the hydrogen lines consist primarily of two close components. Also, the other simple so-called *one-electron spectra*, such as those of Li, Na, K, Rb, and Cs, are doublets with appreciable separations that are easily measured by simple prism or grating spectrographs.

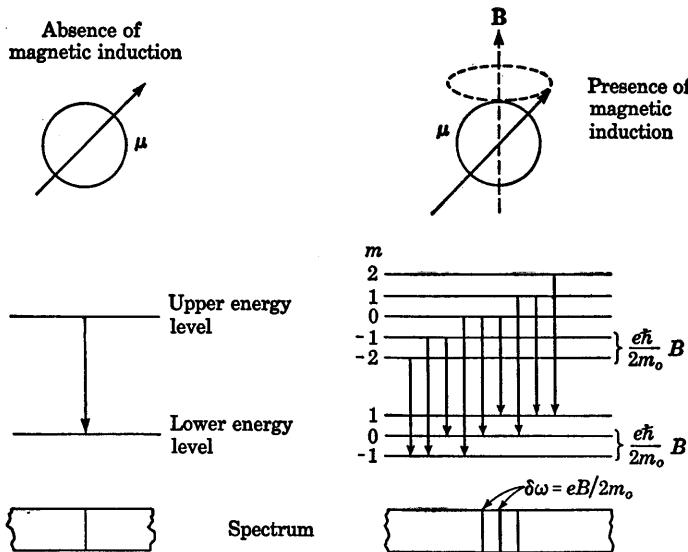


FIG. 4.1. Effect of a magnetic induction upon the energy levels when the magnetic moment has the classical ratio γ_c to the angular momentum.

The classical theory of the magnetic effect of elementary charged particles in rotational motion leads to a further contradiction with experiment in the case of the *Zeeman effect*. This effect is the splitting of spectrum lines into several components when the radiating atoms are in a region of magnetic induction. It is clear from the generality of the expression connecting angular momentum and magnetic moment [Eq. (2.20)] that any type of rotational motion of atomic constituents in the absence of magnetic induction will be modified only to the extent that all of the separate components as well as the system as a whole will process with an angular velocity $\omega' = -(e/2m_0)B$ when in the presence of a magnetic induction B (see Fig. 4.1). Thus any system described by the Schrödinger equation of the preceding chapter will as a whole acquire a precessional angular momentum, say L' , and will have components of angular momentum and magnetic moment along B of $m\hbar$ and $m(e\hbar/2m_0)$, respectively, where m is zero or a positive or negative integer.

The magnetic energy of interaction between the atom and the magnetic induction is $m(e\hbar/2m_0)B$. As the selection rule is $\Delta m = 0, \pm 1$, every spectrum line normally of angular frequency ω would be accompanied by a neighbor on either side at a frequency separation of $eB/2m_0$ in a region of magnetic induction B . Such a pattern is called the *normal Lorentz triplet*, and it is the only one predicted by the foregoing theory. The observational facts, however, are much more complicated. They can be described by saying that each term is split by an amount depending on B into a number of components determined by the nature of the term; that is, the number of electrons of the various types s ($l = 0$), p ($l = 1$), etc., contributing to it. The previous selection rule, $\Delta m = 0, \pm 1$, can then be applied to account for the multiplicity of lines that results. No classical explanation is forthcoming for such behavior.

In order to account for the discrepancies with classical theory that have been outlined above, a major, constructive ideological step was required. This was taken by Uhlenbeck and Goudsmit in 1926 when they proposed that in addition to ordinary orbital or revolving motion, elementary particles exhibit a second type of angular momentum which is one of their intrinsic properties (Sec. 2.6). This is characterized by observable angular-momentum components of $\pm \hbar/2$ along specified axes and a ratio of magnetic moment to this angular momentum of approximately *twice* the classical value. This type of motion has been designated as *spin*. The first of the postulated properties is in contradiction to the discussion in Sec. 3.4 since m takes the values $\pm \frac{1}{2}$, rather than plus or minus an integer. Consequently the description of the axis of spin in terms of the Eulerian angles ϕ and θ is not permissible. Because of the second of the postulated properties, the magnetic moment of a spinning electron is about the same as the magnetic moment associated with orbital motion in the ground state of hydrogen. The consequences of these hypotheses are such that the entire pattern of atomic energy levels and their behavior in the presence of magnetic induction can be qualitatively accounted for. The only apparent obstacle to the achievement of precise quantitative agreement even in the most complicated instances is our inability to carry through completely the mathematical processes required to solve the equations. In principle the resulting differential equations are entirely satisfactory at nonrelativistic energies, and those instances which can be handled mathematically lead to agreement between observation and the prediction of the theory.

The hypothesis of spin as stated above is adequate to describe the general pattern of atomic energy levels and their behavior in a region of magnetic induction. Since there are but two possible orientations of the spin moment with respect to the orbital moment, the fine structures of hydrogen energy levels or terms is double rather than triple as the

classical theory would predict. As has been seen, this is in agreement with experiment. Also, the spin vector tends to precess twice as rapidly under the influence of a magnetic induction as does the orbital vector, since γ of Eq. (2.22) is about twice as great for spin as for orbital motion. What actually happens to an atom under these conflicting precessional tendencies clearly depends on the relative magnitudes of the interaction energies involved and can be worked out quite easily if certain simplifying assumptions are made.

The following discussion, which applies even if the electron structure is a more complicated one than that of hydrogen, is illustrative of the limiting simple cases. It is assumed that if several electrons are present the energy of their spin-orbit interactions is small compared with the Coulomb energy of the configuration, just as the principal energy term for a single electron is the Coulomb term only. This situation is of practical interest as it is found empirically to represent the observations for those atoms for which the outer shell is relatively unpopulated with electrons (Sec. 4.5). It is known as *Russell-Saunders coupling*. The energy of interaction between the spin and orbital motion of the electrons is assumed to predominate over the energy of interaction of these vectors separately with \mathbf{B} , so that the vectors \mathbf{L} and \mathbf{S} precess about a constant resultant \mathbf{J} . \mathbf{J} then precesses slowly about \mathbf{B} , since \mathbf{B} and the energy of interaction of \mathbf{J} with it are small. This approximation corresponds to what is known as the *initial Zeeman effect*.

One is interested in the effective γ , or magnetic moment-angular momentum ratio, for the term since the energy levels are split to the first approximation by an amount proportional to this quantity times the magnetic induction \mathbf{B} . Thus if \mathbf{J} is the total angular momentum of the system, that is, $\mathbf{J}^2 = \hbar^2 j(j+1)$, \mathbf{J} can have $2j+1$ possible projections or components, $j\hbar, (j-1)\hbar, \dots, -(j-1)\hbar, -j\hbar$, along \mathbf{B} . The energies of these are $j\gamma\hbar B, (j-1)\gamma\hbar B$, etc., or writing $g = 2m_0\gamma/e$ and m_j for the values of $j, (j-1), \dots, -(j-1), -j$,

$$\Delta E = g m_j \frac{e\hbar}{2m_0} B \quad (4.1)$$

where g is called the Landé splitting factor.

The frequencies of the spectrum lines are then shifted by the amounts

$$\Delta\omega = \frac{\Delta E_a - \Delta E_b}{\hbar} = (g_a m_{ja} - g_b m_{jb}) \frac{eB}{2m_0}$$

where $m_{ja} - m_{jb} = 0, \pm 1$. The \mathbf{L} and \mathbf{S} vectors are the constant orbital and spin angular-momentum vectors, respectively, which precess

about the constant total angular momentum \mathbf{J} as indicated in Fig. 4.2. The g values can be calculated by reference to the accompanying diagram. The number j is an even or odd multiple of $\frac{1}{2}$ depending on whether there is an even or an odd number of electrons contributing to \mathbf{S} . By hypothesis, \mathbf{L}^2 , \mathbf{S}^2 , and \mathbf{J}^2 are all constant quantities and the component of magnetic moment perpendicular to \mathbf{J} will average out to zero during the precession of \mathbf{L} and \mathbf{S} about \mathbf{J} . Thus it is necessary to know only the component μ_J to find γ which is μ_J/J . On the average,

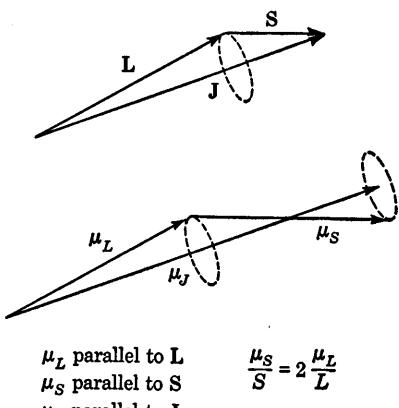


FIG. 4.2. Initial-Zeeman-effect vectors.

From the geometry,

$$\cos \hat{\mathbf{L}}\hat{\mathbf{J}} = \frac{J^2 + L^2 - S^2}{2JL}, \quad \cos \hat{\mathbf{S}}\hat{\mathbf{J}} = \frac{J^2 - L^2 + S^2}{2JS}$$

so

$$g = 1 + \frac{J^2 - L^2 + S^2}{2J^2} = 1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)} \quad (4.2)$$

These results assume a small \mathbf{B} which does not decouple the spin-orbit interaction so that \mathbf{L} and \mathbf{S} form the total angular momentum \mathbf{J} . At the other extreme of large \mathbf{B} , the energy of interaction between the orbital moment and \mathbf{B} and between the spin moment and \mathbf{B} can be much greater than that between \mathbf{L} and \mathbf{S} . In this extreme, \mathbf{L} and \mathbf{S} precess independently about \mathbf{B} as in Fig. 4.3, and the energy of interaction with the magnetic induction is simply

$$\Delta E = (2m_s + m_l) \frac{e\hbar}{2m_0} B \quad (4.3)$$

The small energy correction representing the residual interaction between the components m_s and m_l is here neglected. The selection rules are those derived from classical theory, namely, the change in orbital m_l is 0 or ± 1 and $\Delta m_s = 0$ for electric dipole radiation. Therefore, the transition energies are shifted by $\Delta E = (0, \pm 1) \frac{e\hbar}{2m_0} B$ which corresponds

to the normal Lorentz triplet of classical theory. This extreme of the Zeeman effect for large \mathbf{B} is called the *Paschen-Back effect*.

The component lines in a Zeeman multiplet are polarized in the senses that one would expect intuitively on the basis of the discussion of the radiation of a circulating electron in Sec. 1.8. A positive change in angular-momentum component along \mathbf{B} ($\Delta m_l = 1$) corresponds to the establishment of a greater electron circulation in such a sense as to increase the energy of the magnetic moment, the frequency emitted is greater, and if observed in the direction of \mathbf{B} the radiation is circularly polarized in a clockwise sense. If observed at right angles to \mathbf{B} , the radiation is plane-polarized at right angles to \mathbf{B} . Conversely a negative change in angular-momentum component along \mathbf{B} ($\Delta m_l = -1$) leads to a lower frequency and counterclockwise polarization when observed along \mathbf{B} and the same sense of plane polarization when observed normal to \mathbf{B} . No change in the angular-momentum component along \mathbf{B} ($\Delta m_l = 0$) corresponds to a linear type of dipole oscillation along \mathbf{B} ; the energy is unchanged by the presence of the magnetic induction, there is zero intensity in the direction of the dipole axis along \mathbf{B} , and the radiation is plane-polarized in the plane of \mathbf{B} when observed normal to the magnetic induction.

It should be emphasized that the foregoing discussion is not general in the sense that the dependence of energy levels on magnetic induction for intermediate values of \mathbf{B} is not described. Equations (4.1) and (4.3) apply only when ΔE is either very small or very great in comparison with the energy of interaction of the component atomic magnetic vectors among themselves. Essentially, these equations describe classical cases for which γ or g take certain simple, constant values. A quantum-mechanical description involves a deeper understanding of the phenomena. The following tables and diagrams illustrate the two extremes. The nomenclature is that devised to describe spectroscopic terms. It is a simple extension of the assignment of letters to the angular momentum values of Sec. 3.4. By a *configuration* is meant the description of the number of electrons of different angular-momentum types, s ($l = 0$), p ($l = 1$), d ($l = 2$), etc., in the different total quantum states, $n = 1, 2, 3, \dots$. A preceding integer gives the total quantum number n , and a superscript indicates the number of electrons of the type with that total quantum number, that is, $1s$, $1s^2$, $2p^2$, etc. By a *term* is meant the description in terms of orbital, spin, and total angular momentum of the structure to which the configuration may give rise. In the Russell-Saunders coupling

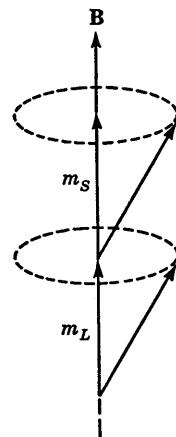


FIG. 4.3. Complete Paschen-Back-effect vectors.

scheme, large letters are used by analogy with the small ones for indicating total orbital angular momentum. Superscripts corresponding to the multiplicity ($2S + 1$) indicate the total spin angular momentum, and subscripts indicate the total angular momentum J , the magnitude of the various possible combinations of L and S . In other more general types of coupling where the electrostatic forces do not clearly predominate, the magnetic interactions between the electron orbital and spin moments do not permit separate conservation of these two types of angular momentum. The total angular momentum J is alone conserved, and no significance can be attached to an L or an S .

4.2. Consequences of a Relativistic Formulation. The spin concept was introduced in an ad hoc manner in the preceding section, and, though

TABLE 4.1

	Configuration	Terms	
1 electron	s^1	$^2S_{\frac{1}{2}}$	
	p^1	$^2P_{\frac{1}{2}}$	$^2P_{\frac{3}{2}}$
	d^1	$^2D_{\frac{3}{2}}$	$^2D_{\frac{5}{2}}$
	.	.	.
	.	.	.
	ss	1S_0	3S_1
	sp	1P_1	$^3P_{0,1,2}$
	sd	1D_2	$^3D_{1,2,3}$
	.	.	.
	.	.	.

TABLE 4.2. ONE-ELECTRON ZEEMAN-TERM PATTERNS (illustrated in Fig. 4.4)

Term	Initial Zeeman effect			Extreme Paschen-Back effect		
	m_j	g	$m_j g$	m_l	m_s	$m_l + 2m_s$
$^2S_{\frac{1}{2}}$	$\frac{1}{2}$	2	1	0	$\frac{1}{2}$	1
	$-\frac{1}{2}$	2	-1	0	$-\frac{1}{2}$	-1
$^2P_{\frac{1}{2}}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{3}$	1	$\frac{1}{2}$	2
	$-\frac{1}{2}$	$\frac{2}{3}$	$-\frac{1}{3}$	0	$\frac{1}{2}$	1
$^2P_{\frac{3}{2}}$	$\frac{3}{2}$	$\frac{4}{3}$	2	1	$-\frac{1}{2}$	0
	$\frac{1}{2}$	$\frac{4}{3}$	$\frac{2}{3}$	0	$-\frac{1}{2}$	-1
	$-\frac{1}{2}$	$\frac{4}{3}$	$-\frac{2}{3}$	-1	$-\frac{1}{2}$	-2
	$-\frac{3}{2}$	$\frac{4}{3}$	-2	.	.	.

and so forth for D , F , G , etc., terms

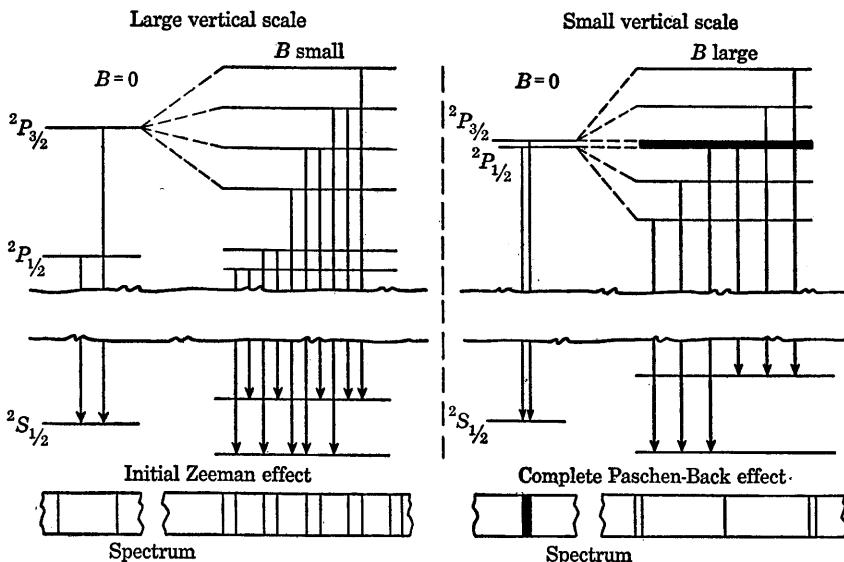


FIG. 4.4. Illustrative one-electron Zeeman effect, extreme cases.

its utility was seen in its resolution of the two primary spectroscopic difficulties that the classical theory presented, it does not fit naturally into the treatment of the preceding chapter. However, if the Schroedinger equation is written in a form proposed by Dirac¹ which is concordant with the special theory of relativity, it is found that the attribute of spin is a necessary consequence. In addition, the theory predicts the existence of *positrons*, particles discovered by Anderson² in 1932, which are otherwise similar to electrons but have a positive charge. Finally, the theory describes the proper energetic relations between spin and orbital moments and between these moments and a magnetic induction \mathbf{B} . The theory of relativity simplified and unified classical mechanics and electricity when applied to the equations of Newton and Maxwell. In much the same way, Dirac's theory is found to simplify the equation of Schrödinger and incorporate the concept of spin naturally into the resulting structure.

The equations of Sec. 3.4 would not retain their form under a Lorentz transformation [Eq. (1.13)]. Also the classical equation (1.16), which is satisfactory relativistically, is not suitable for conversion to a wave equation by the substitutions $\mathbf{E} \rightarrow +i\hbar \partial/\partial t$ and $\mathbf{p} \rightarrow -i\hbar \nabla$. The reasons for this are: (a) the resulting position-probability density function is not

¹ P. A. M. Dirac, *Proc. Roy. Soc. (London)*, **117**, 610 (1928).

² C. D. Anderson, *Phys. Rev.*, **43**, 481 (1933).

always positive, leading to meaningless results, and (b) the equation contains the second derivative with respect to t rather than the first. As mentioned in connection with Eq. (3.6), the equation describing the physical situation should contain only the first derivative of ψ with respect to t . If ψ is the value of the wave function in some instant, its value a very short time δt afterwards should be $\psi + (\partial\psi/\partial t)\delta t$. Hence, $\partial\psi/\partial t$, not $\partial^2\psi/\partial t^2$, should be used in any equation purporting to describe the sequence of events. Dirac's approach was to assume an equation linear in $\partial/\partial t$ and $\partial/\partial x$ so that the spatial and temporal variables are treated equivalently as the relativistic theory requires. Dirac then determined the coefficients in such a way that the equation agrees with the classical equations in terms of the energy and momentum concepts. Thus the proper energy equation for a free particle of rest mass m_0 is assumed to be

$$H\psi = (c\mathbf{a} \cdot \mathbf{p} + \beta m_0 c^2)\psi = i\hbar \frac{\partial \psi}{\partial t} = E\psi \quad (4.4)$$

or

$$\left(i\hbar \frac{\partial}{\partial t} + i\hbar c\mathbf{a} \cdot \nabla - \beta m_0 c^2 \right) \psi = 0 \quad (4.4')$$

The vector \mathbf{a} and the scalar β are to be determined by squaring Eq. (4.4) divided by c , and comparing with Eq. (1.16). This yields

$$\begin{aligned} \frac{E^2}{c^2} \psi &= [a_x^2 p_x^2 + a_y^2 p_y^2 + a_z^2 p_z^2 + \beta^2 m_0^2 c^2 \\ &\quad + (a_x a_y + a_y a_x) p_x p_y + (a_y a_z + a_z a_y) p_y p_z \\ &\quad + (a_z a_x + a_x a_z) p_x p_z + (a_x \beta + \beta a_x) p_x m_0 c \\ &\quad + (a_y \beta + \beta a_y) p_y m_0 c + (a_z \beta + \beta a_z) p_z m_0 c] \psi \end{aligned}$$

To agree with

$$\left(\frac{E}{c} \right)^2 = p^2 + m_0^2 c^2 \quad (1.16)$$

the following relations must hold:

$$\begin{aligned} a_x^2 &= a_y^2 = a_z^2 = \beta^2 = 1 \\ a_x a_y + a_y a_x &= a_y a_z + a_z a_y = a_z a_x + a_x a_z = 0 \\ a_x \beta + \beta a_x &= a_y \beta + \beta a_y = a_z \beta + \beta a_z = 0 \end{aligned} \quad (4.5)$$

It is clear that there are no ordinary algebraic numbers that satisfy Eqs. (4.5) any more than such numbers could be found to satisfy Eqs. (3.17), but certain entities other than differential operators which are familiar in other fields of physics do have the necessary properties. These are so-called *matrices*, which are familiar, though possibly not under that

name, in almost all branches of classical physics.¹ They are most commonly encountered in dealing with sets of linear equations such as transformations from one set of orthogonal axes to another. Consider a set of potential differences that are related to a set of currents by the impedances of a network, or a set of strains that are related to a set of stresses by the properties of an elastic solid, or a set of vector components related to another set by a rotation. The relationship in all these instances is a series of linear equations:

$$x_i = \sum_j a_{ij} x'_j$$

where the set of a_{ij} 's, called a matrix, characterizes the network, elastic solid, or rotation. The set of a_{ij} 's may be most conveniently thought of as constituting a square array in which the entries occupy the relative positions of the coefficients in a set of homogeneous linear equations; the particular coefficient a_{ij} occurring at the intersection of the i th row and j th column.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

If one rotation, say, specified by the a 's, is followed by another specified by b 's,

$$x'_j = \sum_k b_{jk} x''_k$$

the result of the two rotations can be written as $x_i = \sum_k c_{ik} x''_k$, where $c_{ik} = \sum_j a_{ij} b_{jk}$ as can be seen by making the substitutions. If these rotations are performed in the reverse order, that is, $x_i = \sum_j b_{ij} x'_j$ followed by $x'_j = \sum_k a_{jk} x''_k$, the result is $x_i = \sum_k d_{ik} x''_k$, where $d_{ik} = \sum_j b_{ij} a_{jk}$, and it can be seen that in general $c_{ik} \neq d_{ik}$. These matrix equations may be written in the following symbolic form:

$$x = Ax', \quad x' = Bx'', \quad x = Cx'', \quad C = AB$$

Sans-serif type is used to designate matrices as well as differential operators, as both these exhibit the noncommuting property. The x 's are one-column matrices obeying the same law of multiplication as the square

¹ For an outline of the mathematical properties of matrices see for instance H. Margenau and G. M. Murphy, "Mathematics of Physics and Chemistry," D. Van Nostrand Company, Inc., New York, 1948.

matrices A and B. From the law of multiplication it is seen that

$$AB \neq BA$$

which is just the type of relation for quantities that are to obey Eqs. (3.17) or (4.5). Thus matrices may be thought of as characterizing the quantum-mechanical quantities, just as matrices characterize electric networks or elastic or piezoelectric bodies.

The entries in the array constituting the matrix, which are known as the *matrix elements*, can be derived from the conditions that they must satisfy, such as Eqs. (3.17) and (4.5), as may be seen by consulting the references on quantum mechanics. Here their forms will simply be given, and the reader can verify by use of the multiplication law that they satisfy the defining equations. The matrices representing the L 's are square arrays consisting of $(2l + 1)$ rows and columns. The elements of L_z are zero everywhere except along the diagonal from the upper left to the lower right, known as the *main diagonal*, which has the values $m_l\hbar$, where m_l takes the successive values $l, (l - 1), (l - 2), \dots, -l$ from the upper-left corner to the lower-right corner. If a system of homogeneous linear equations representing a transformation involving $(2l + 1)$ variables, which may be written as Y_{l,m_l} and thought of as representing the components of a vector in a space of $(2l + 1)$ dimensions, is written out using the matrix L_z as the array of coefficients, it is seen that these are very simple equations involving only the analogous component of the vector in each representation. The equations are

$$L_z Y_{l,m_l} = m_l \hbar Y_{l,m_l}$$

where m_l takes the series of values from $+l$ to $-l$. This of course corresponds exactly to Eq. (3.18). A constant such as \hbar may be thought of as a matrix, the elements of which are zero everywhere except along the main diagonal where they have the value of the constant. The zero matrix has all zero elements. The elements of the matrix L_x are zero everywhere except along the diagonals immediately above and immediately below the main one. Along the upper diagonal the elements are $(\hbar/2)[l(l + 1) - m_l(m_l - 1)]^{\frac{1}{2}}$ or $(\hbar/2)[(l - m_l + 1)(l + m_l)]^{\frac{1}{2}}$, where m_l takes the successive values $l, (l - 1), (l - 2), \dots, -(l - 1)$. The lower-diagonal elements are such as to make the matrix symmetrical about the main diagonal. The elements of L_y are the same as for L_x except that they are multiplied by $\sqrt{-1}$ and the signs of the elements in the upper diagonal are changed from plus to minus. If these matrices are used as the pattern for a linear transformation, the equations are not simple ones relating analogous components only, as in the case of L_z . Either one could of course be made so by a suitable rotation at the expense

of destroying the simplicity for \mathbf{l}_z . Finally, the elements of

$$\mathbf{l}^2 = \mathbf{l}_x^2 + \mathbf{l}_y^2 + \mathbf{l}_z^2$$

can be formed from the components, and they are seen to differ from zero only along the main diagonal. Here they are all the same and equal to $\hbar^2 l(l+1)$, so \mathbf{l}^2 is a constant. As an example, for $l = 1$,

$$\begin{aligned}\mathbf{l}_x &= \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, & \mathbf{l}_y &= \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix} \\ \mathbf{l}_z &= \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, & \mathbf{l}^2 &= 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 2\hbar^2\end{aligned}$$

As an exercise these should be substituted in Eq. (3.17) to verify that the equations are satisfied by these matrices.

Returning to the matrices for \mathbf{a} and β which satisfy Eqs. (4.5), we find these to be of four rows and four columns but quite closely related in form to the \mathbf{l} matrices for $l = \frac{1}{2}$. These matrices satisfying Eqs. (4.5) are

$$\begin{aligned}\mathbf{a}_x &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, & \mathbf{a}_y &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \\ \mathbf{a}_z &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, & \beta &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\end{aligned}$$

The fact that these are solutions should be verified by direct substitution using the matrix multiplication rule $c_{ik} = \sum_j a_{ij} b_{jk}$. The connection between the \mathbf{a} matrices and those for \mathbf{l} ($l = \frac{1}{2}$) is seen by noting that the two-row and two-column submatrices in the upper right and lower left corners are the same as the corresponding two-row and two-column \mathbf{l} matrices. Three component matrices of a new matrix σ , known as the *Pauli spin matrix*, are defined as

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Hence, formally,

$$\mathbf{l}(l = \frac{1}{2}) = \frac{\hbar}{2} \sigma$$

α may be written in abbreviated form as a two-row, two-column matrix:

$$\alpha = \begin{pmatrix} 0 & \delta \\ \delta & 0 \end{pmatrix}$$

or

$$\alpha_x = \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}, \quad \alpha_y = \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix}, \quad \alpha_z = \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}$$

The two-row, two-column matrix δ may be written as a four-row, four-column matrix by repeating δ along the main diagonal. The justification for extending the number of rows and columns, known as the *order* of the matrix, in this way is evident upon looking at the pattern of coefficients of linear transformation equations. Thus

$$\delta = \begin{pmatrix} \delta & 0 \\ 0 & \delta \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

etc. The α and δ matrices have certain properties in addition to those of Eqs. (4.5).

$$\sigma_x \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i \sigma_z = -\sigma_y \sigma_x$$

Or in general

$$\begin{aligned} \sigma_x \sigma_y - \sigma_y \sigma_x &= 2i \sigma_z \\ \sigma_y \sigma_z - \sigma_z \sigma_y &= 2i \sigma_x \\ \sigma_z \sigma_x - \sigma_x \sigma_z &= 2i \sigma_y \end{aligned} \tag{4.6}$$

It is evident that if α and β are matrices of four rows and columns, Eq. (4.4) must be an equation in four component ψ 's, or Eq. (4.4) represents four linear equations instead of one. In matrix notation ψ would be written as a one-column matrix, and the *conjugate matrix*, which is the one in which the columns are written for rows and the rows for columns with a change in sign of an imaginary part, as a row matrix:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \quad \psi^* = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$$

In accordance with the law of matrix multiplication,

$$\psi^* \psi = \psi_1 \psi_1^* + \psi_2 \psi_2^* + \psi_3 \psi_3^* + \psi_4 \psi_4^*$$

plays the same role in the theory as the previous $\psi \psi^*$ where the ψ 's were scalar rather than matrix quantities. For the purpose of discussing

atomic structure it is unnecessary to develop the complete relativistic theory in terms of these matrices, and only three of its consequences will be demonstrated: (a) the conservation of total angular momentum, (b) the form of the interaction energy with \mathbf{B} , and (c) the form of the spin-orbit interaction energy.

Application to the Hydrogen Atom. The generalization of Eq. (4.4) to describe the electron-proton combination composing a hydrogen atom is fraught with great difficulty because of our inadequate grasp of the properties characterizing elementary particles, and no completely satisfactory formulation has as yet been proposed. But as a heuristic approximation it may be tentatively assumed that the effect of the proton is represented by the inclusion in Eq. (4.4) of a spherically symmetric potential V due to the proton. Thus as a starting point it will be assumed that

$$H\psi = (E + V)\psi = (c\mathbf{a} \cdot \mathbf{p} + \beta m_0 c^2 + V)\psi \quad (4.7)$$

It should be noted that since it has been shown only that E^2 is equivalent to $(c\mathbf{a} \cdot \mathbf{p} + \beta m_0 c^2)^2$ there is an ambiguity in sign which has been arbitrarily resolved as positive in Eq. (4.7). This choice is not of any particular significance in the nonrelativistic discussions that follow; the possible negative energy states are of importance in the theory of the positron for which the references should be consulted. Since $E = i\hbar \partial/\partial t$, by the Schrödinger convention, any matrix that commutes with H or E must be independent of t . This may be seen by writing the expression $HA - AH$ for an arbitrary matrix A . Since $E\psi = i\hbar \partial\psi/\partial t$,

$$\begin{aligned} HA\psi &= i\hbar \frac{\partial A\psi}{\partial t} + VA\psi = i\hbar \frac{\partial A\psi}{\partial t} + i\hbar A \frac{\partial \psi}{\partial t} + VA\psi \\ &= i\hbar \psi \frac{\partial A}{\partial t} + AH\psi \end{aligned}$$

or

$$HA\psi - AH\psi = i\hbar \psi \frac{\partial A}{\partial t}$$

Thus if $HA - AH = 0$, A does not depend on t , and the matrix A is a constant of the motion. It can be shown that \mathbf{L} does not satisfy this condition, but a matrix $\mathbf{J} = \mathbf{L} + \mathbf{S}$ does where $\mathbf{S} = \frac{1}{2}\hbar\boldsymbol{\sigma}$. By definition $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ where $\mathbf{p} = -i\hbar\nabla$. Thus, since V is a scalar and spherically symmetrical function of \mathbf{r} , and \mathbf{a} and β do not contain the variables (x, y, z) , the x component of $\partial\mathbf{L}/\partial t$ is seen by the use of Eq. (4.7) to be

$$-i\hbar \frac{\partial L_x}{\partial t} = L_x H - H L_x = \hbar^2 c \left(a_y \frac{\partial}{\partial z} - a_z \frac{\partial}{\partial y} \right)$$

also

$$-i\hbar \frac{\partial S_x}{\partial t} = S_x H - HS_x$$

and since the σ , a , and β matrices do not commute,

$$\begin{aligned} -i\hbar \frac{\partial S_x}{\partial t} = & -\frac{i}{2} \hbar^2 c \left[(\sigma_x a_x - a_x \sigma_x) \frac{\partial}{\partial x} + (\sigma_x a_y - a_y \sigma_x) \frac{\partial}{\partial y} \right. \\ & \left. + (\sigma_x a_z - a_z \sigma_x) \frac{\partial}{\partial z} \right] + \frac{\hbar}{2} (\sigma_x \beta - \beta \sigma_x) m_0 c^2 \end{aligned}$$

Using the definition of the σ and a components, this is seen to be equal to

$$\begin{aligned} -i\hbar \frac{\partial S_x}{\partial t} = & -\frac{i}{2} \hbar^2 c \left(2ia_z \frac{\partial}{\partial y} - 2ia_y \frac{\partial}{\partial z} \right) \\ = & \hbar^2 c \left(a_z \frac{\partial}{\partial y} - a_y \frac{\partial}{\partial z} \right) \end{aligned}$$

Comparing this with $\partial L_x / \partial t$ above, it is seen that $\partial(L_x + S_x)/\partial t = 0$. Since similar expressions hold for the y and z components, $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is constant, as the total angular momentum must be on any consistent description.

In order to consider Eq. (4.4) in the presence of a magnetic induction, $\mathbf{p} - e\mathbf{A}$ must be used for \mathbf{p} where \mathbf{A} is the magnetic vector potential. Equation (4.7) thus becomes

$$H\Psi = [V + c\mathbf{a} \cdot (\mathbf{p} - e\mathbf{A}) + \beta m_0 c^2]\Psi \quad (4.8)$$

To see what this implies for the energy in terms of magnetic moments and the magnetic induction, the square of the kinetic energy from Eq. (4.8) must be compared with that from Eq. (1.16) and the nonrelativistic Hamiltonian written to the first approximation in terms of v^2/c^2 . From Eq. (4.8),

$$E^2 = (H - V)^2 = [c\mathbf{a} \cdot (\mathbf{p} - e\mathbf{A}) + \beta m_0 c^2]^2$$

From Sec. 1.6,

$$E = mc^2 = m_0 c^2 + \frac{1}{2} m_0 v^2 + \dots = m_0 c^2 + E'$$

where E' is the nonrelativistic kinetic energy. Neglecting higher powers of v^2/c^2 , E^2 is given by $E^2 = (m_0 c^2)^2 + 2m_0 c^2 E'$. Equating these two expressions for E^2 and E^2 , as these quantities must be equivalent in the nonrelativistic approximation,

$$2m_0 c^2 E' = c^2 \{ [\mathbf{a} \cdot (\mathbf{p} - e\mathbf{A})][\mathbf{a} \cdot (\mathbf{p} - e\mathbf{A})] \}$$

Since $\beta^2 = 1$ and $\mathbf{a}\beta = -\beta\mathbf{a}$. Then the nonrelativistic Hamiltonian is

$H' = E' + V$, and

$$H'\psi = \left(V + \frac{1}{2m_0} \{ [\mathbf{a} \cdot (\mathbf{p} - e\mathbf{A})] [\mathbf{a} \cdot (\mathbf{p} - e\mathbf{A})] \} \right) \psi$$

so the product in the braces must be calculated to obtain the nonrelativistic energy equation. The \mathbf{a} 's do not commute with each other nor do the \mathbf{p} 's commute with the vector potential \mathbf{A} . The expression in the braces is then

$$[\alpha_x(p - eA)_x + \alpha_y(p - eA)_y + \alpha_z(p - eA)_z]^2$$

and, since $\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = 1$, using Eqs. (4.5) and noting that $\alpha_x\alpha_y = i\sigma_z$, etc.,

$$\{ \} = (\mathbf{p} - e\mathbf{A})^2 + i\boldsymbol{\sigma} \cdot (\mathbf{p} - e\mathbf{A}) \times (\mathbf{p} - e\mathbf{A})$$

Forming the vector product and recalling that $\mathbf{p} = -i\hbar\nabla$,

$$(\mathbf{p} - e\mathbf{A}) \times (\mathbf{p} - e\mathbf{A}) = i\hbar e \nabla \times \mathbf{A} = i\hbar e \mathbf{B}$$

and, finally,

$$H'\psi = \left[V + \frac{1}{2m_0} (\mathbf{p} - e\mathbf{A})^2 - \frac{e\hbar}{2m_0} \boldsymbol{\sigma} \cdot \mathbf{B} \right] \psi \quad (4.9)$$

The first two terms in the parentheses are those of the ordinary Schroedinger theory extended to include the presence of magnetic induction by replacing \mathbf{p} by $\mathbf{p} - e\mathbf{A}$. The last term, which comes from the matrix nature of Eq. (4.8), is the spin-moment magnetic energy and is seen to be of the form of the energy due to a magnetic moment $e\hbar/2m_0$ oriented either parallel or antiparallel to the magnetic induction. The ratio of the spin magnetic moment $(e\hbar/2m_0)\boldsymbol{\sigma}$ to the spin angular momentum $\mathbf{S} = \frac{1}{2}\hbar\boldsymbol{\sigma}$ is e/m_0 , or twice the classical value. Thus the spin properties necessary to account for the observed spectroscopic phenomena are consequences of a relativistic formulation of the theory and do not have to be introduced artificially as they were in the historical development.

Finally, it may be seen that the relativistic equation leads to a proper description of the interaction energy of the spin and orbital magnetic moments of a charged particle in a spherically symmetric electric potential $V(r)$. In the nonrelativistic discussion of the preceding paragraph it was seen that $E \cong E' + m_0c^2$. In consequence $H \cong H' + m_0c^2$ and Eq. (4.7) may be written in this approximation as

$$(H' + m_0c^2)\psi = [V(r) + c\boldsymbol{\sigma} \cdot \mathbf{p} + \beta m_0c^2]\psi$$

In the following discussion one cannot avoid using ψ 's as one-column matrices, but on writing out the four equations it will be seen that they

can be grouped two-and-two so that if ψ' represents $\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ and ψ'' represents $\begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}$ the four equations can be written as two:

$$\begin{aligned} (\mathcal{H}' + m_0c^2)\psi' &= [V(r) + m_0c^2]\psi' + c\mathbf{\sigma} \cdot \mathbf{p}\psi'' \\ (\mathcal{H}' + m_0c^2)\psi'' &= c\mathbf{\sigma} \cdot \mathbf{p}\psi' + [V(r) - m_0c^2]\psi'' \end{aligned}$$

Collecting terms and dividing through by $2m_0c^2$ the second equation is

$$\left[1 + \frac{\mathcal{H}' - V(r)}{2m_0c^2} \right] \psi'' = \frac{1}{2m_0c} \mathbf{\sigma} \cdot \mathbf{p} \psi'$$

In the nonrelativistic limit here under consideration the second term in the bracket on the left is small.¹ Assuming the ψ 's to be eigenfunctions of the operator \mathcal{H}' with the energy eigenvalues E_n , the reciprocal of the parenthesis may be expanded retaining only first powers of the small quantity, yielding

$$\psi'' = \frac{1}{2m_0c} \left[1 - \frac{E_n - V(r)}{2m_0c^2} \right] \mathbf{\sigma} \cdot \mathbf{p} \psi'$$

Substituting this value of ψ'' in the first of the two equations,

$$\mathcal{H}'\psi' = \left\{ V(r) + \frac{1}{2m_0} \mathbf{\sigma} \cdot \mathbf{p} \left[1 - \frac{E_n - V(r)}{2m_0c^2} \right] \mathbf{\sigma} \cdot \mathbf{p} \right\} \psi'$$

Writing out $\mathbf{\sigma} \cdot \mathbf{p}$ as $-i\hbar(\sigma_x \partial/\partial x + \sigma_y \partial/\partial y + \sigma_z \partial/\partial z)$, it is seen that $(\mathbf{\sigma} \cdot \mathbf{p})(\mathbf{\sigma} \cdot \mathbf{p}) = \mathbf{p}^2$. Also

$$(\mathbf{\sigma} \cdot \mathbf{p})V(r)(\mathbf{\sigma} \cdot \mathbf{p}) = V(r)\mathbf{p}^2 - i\hbar \frac{1}{r} \frac{\partial V(r)}{\partial r} (\mathbf{\sigma} \cdot \mathbf{r})(\mathbf{\sigma} \cdot \mathbf{p})$$

and

$$(\mathbf{\sigma} \cdot \mathbf{r})(\mathbf{\sigma} \cdot \mathbf{p}) = \mathbf{r} \cdot \mathbf{p} + i(\mathbf{\sigma} \cdot \mathbf{r} \times \mathbf{p})$$

Thus the equation becomes

$$\begin{aligned} \mathcal{H}'\psi' &= \left\{ V(r) + \frac{\mathbf{p}^2}{2m_0} \left[1 - \frac{E_n - V(r)}{2m_0c^2} \right] - \frac{i\hbar}{4m_0^2c^2} \frac{1}{r} \frac{\partial V(r)}{\partial r} \mathbf{r} \cdot \mathbf{p} \right. \\ &\quad \left. + \frac{\hbar}{4m_0^2c^2} \frac{1}{r} \frac{\partial V(r)}{\partial r} (\mathbf{\sigma} \cdot \mathbf{r} \times \mathbf{p}) \right\} \psi' \end{aligned}$$

Since $E_n - V(r)$ is the nonrelativistic kinetic energy $p^2/2m_0$, the first two terms on the right differ from those of the nonrelativistic Schroedinger equation only by a term that is smaller than those of Eq. (3.8) by the factor $(p/2m_0c)^2$. The third term is also a small relativistic one which is difficult to observe experimentally as it may be shown to be

¹ Except for very near the nucleus where $V(r)$ becomes large as $1/r$. This is of importance only for an s term, for, if $l > 0$, ψ'' vanishes for $r = 0$.

finite only for an s term. The fourth term is of course not large either, but its effect is readily evident spectroscopically because the vectors involved can take different relative orientations leading to a directly measurable series of energy differences. Using the definitions $\mathbf{S} = \frac{1}{2}\hbar\mathbf{s}$ and $\mathbf{L} = \mathbf{r} \times \mathbf{p}$,

$$\mathcal{H}'\psi' = \left\{ V(r) + \frac{\mathbf{p}^2}{2m_0} \left[1 - \left(\frac{\mathbf{p}}{2m_0c} \right)^2 \right] - \frac{\hbar^2}{4m_0^2c^2} \frac{\partial V(r)}{\partial r} \frac{\partial}{\partial r} + \frac{1}{2m_0^2c^2} \frac{1}{r} \frac{\partial V(r)}{\partial r} \mathbf{S} \cdot \mathbf{L} \right\} \psi' \quad (4.10)$$

The last term is thus seen to represent the interaction between the spin and orbital moments.

A relativistic calculation of the foregoing type is necessary to derive the spin-orbit interaction correctly, as it has its genesis essentially in relativistic interactions. In the nonrelativistic approximation considered here the equation for ψ'' is not of interest since it is seen from earlier equations in this paragraph that, neglecting terms of the order of $(v/c)^2$, $\psi'' = (\mathbf{s} \cdot \mathbf{p}/2m_0c)\psi'$ and hence ψ'' is smaller than ψ' by the ratio v/c , which justifies its neglect.

In the case of a simple Coulomb-type potential, such as that due to an isolated nucleus of charge Ze , $V(r)$ is given by $-Ze^2/4\pi\epsilon_0 r$ and the spin-orbit interaction term becomes

$$\frac{1}{2m_0^2c^2} \frac{1}{r} \frac{\partial V(r)}{\partial r} \mathbf{S} \cdot \mathbf{L} = \frac{Ze^2}{8\pi\epsilon_0 m_0^2 c^2} \frac{1}{r^3} \mathbf{S} \cdot \mathbf{L}$$

The mean value of $1/r^3$ may be calculated using the definition

$$\left(\frac{1}{r^3} \right)_{av} = \int_0^\infty \psi(1/r^3)\psi^* dv$$

according to Sec. 3.4, and, if hydrogenic wave functions as given in that section are used, the result is

$$\left(\frac{1}{r^3} \right)_{av} = \frac{Z^3 m_e^3 e^6}{(4\pi\epsilon_0)^3 \hbar^6 n^3 l(l + \frac{1}{2})(l + 1)}$$

This expression applies, of course, only if there is a finite orbital angular momentum, that is, l is equal to or greater than 1. If $l = 0$, a reconsideration of the approximation that $[\mathcal{H}' - V(r)]/2m_0c^2$ is small shows that $(1/r^2)_{av}$ remains finite and the contribution of this term to Eq. (4.10) vanishes because $\mathbf{L} \cdot \mathbf{S}$ vanishes. The average value for the interaction energy of \mathbf{L} and \mathbf{S} for a hydrogenlike atom is then

$$U_{LS} = \eta \mathbf{S} \cdot \mathbf{L} \quad (4.11)$$

Neglecting the difference between the reduced mass m_r and m_0 , the constant η may be written more briefly in terms of the Rydberg constant for infinite mass $R_\infty = m_0 e^4 / [4\pi(4\pi\epsilon_0)^2 \hbar^3 c]$ and the fine-structure constant $\alpha = e^2 / 4\pi\epsilon_0 \hbar c$ as

$$\eta = \frac{2\pi R_\infty \alpha^2 c Z^4}{\hbar n^3 l(l + \frac{1}{2})(l + 1)}$$

Equation (4.11) gives a very satisfactory description of the fine structure of the spectra of atomic hydrogen, singly ionized helium, doubly ionized lithium, etc. If more than one electron is present in the atomic structure, the potential $V(r)$ is no longer of the Coulomb type and the wave functions are no longer of the hydrogenic type, so the value given for the constant η is not appropriate to describe the energy separation of the alkali-metal doublets as shown in Table 4.3. However, the general form of the dependence of the energy upon the spin and orbital angular-momentum vectors given by Eq. (4.11) is applicable to many polyatomic atomic structures. It gives the general nature of the multiplicity of the energy levels due to this cause, including the number of component terms and their relative displacements.

A considerable number of the chemical elements, particularly those on the left-hand side of the periodic table, have electron structures of such a type that due to strong electrostatic interactions between the several orbital angular-momentum vectors they tend to combine to yield a constant orbital momentum \mathbf{L} for the structure as a whole. Similarly, strong electrostatic interactions tend to combine the spin vectors of the several electrons into a constant spin angular momentum \mathbf{S} characteristic of the structure. This is the case of Russell-Saunders coupling (Sec. 4.1), and in such a case the entire structure is formally analogous to a single electron in that the total angular momentum, orbital angular momentum, and spin angular momentum all separately remain constant in magnitude in whatever motion the structure undergoes. In consequence of the constancy of each of these vectors separately for a single-electron structure or one in which Russell-Saunders coupling exists, $\mathbf{J}^2 = \hbar^2 j(j + 1)$, $\mathbf{L}^2 = \hbar^2 l(l + 1)$, and $\mathbf{S}^2 = \hbar^2 s(s + 1)$. By the law of triangles,

$$\mathbf{J}^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L} \cdot \mathbf{S}$$

so for a one-electron atom or for Russell-Saunders coupling $\mathbf{L} \cdot \mathbf{S}$ is also constant and Eq. (4.11) may be written

$$\begin{aligned} U_{LS} &= \eta \mathbf{L} \cdot \mathbf{S} = \frac{\eta}{2} (\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \\ &= \eta \frac{\hbar^2}{2} [j(j + 1) - l(l + 1) - s(s + 1)] \end{aligned} \quad (4.11')$$

For a single electron the interaction energy is

$$U_{LS} = \frac{\pi \hbar R_\infty \alpha^2 c Z^4 [j(j+1) - l(l+1) - s(s+1)]}{n^3 l(l+\frac{1}{2})(l+1)}$$

The number of components of a term or energy level arising from a given L and S is the number of total angular-momentum possibilities, or the

TABLE 4.3. REPRESENTATIVE MULTIPLET SEPARATIONS†

Element	Outer electron configuration	Term differences, cm^{-1}	Ratio	
			Experiment	Theory
H	2p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 0.365 (R_H \alpha^2 / 16)$		
He ⁺	2p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 5.84 (16R_{\text{He}} \alpha^2 / 16)$		
Li ⁺⁺	2p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 29.5 (81R_{\text{Li}} \alpha^2 / 16)$		
Li	2p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 0.54$		
Na	3p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 17.20$		
K	4p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 57.72$		
	5p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 18.76$		
	6p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 8.41$		
	7p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 4.48$		
	8p	$^2P_{\frac{3}{2}} = ^2P_{\frac{1}{2}} = 2.70$		
	9p	$^2P_{\frac{3}{2}} - ^2P_{\frac{1}{2}} = 1.74$		
Mg	3s3p	$^3P_2 - ^3P_1 = 40.714$	$^3P_1 - ^3P_0 = 20.058$	1.98
	3p ²	$^3P_2 - ^3P_1 = 40.61$	$^3P_1 - ^3P_0 = 20.56$	1.97
Na	4s4p	$^3P_2 - ^3P_1 = 105.881$	$^3P_1 - ^3P_0 = 52.158$	2.03
Ca	4s3d	$^3D_3 - ^3D_2 = 21.44$	$^3D_2 - ^3D_1 = 13.90$	1.54
	4s4d	$^3D_3 - ^3D_2 = 5.578$	$^3D_2 - ^3D_1 = 3.682$	1.52
	4s5d	$^3D_3 - ^3D_2 = 2.667$	$^3D_2 - ^3D_1 = 1.718$	1.55
	4s6d	$^3D_3 - ^3D_2 = 1.953$	$^3D_2 - ^3D_1 = 1.340$	1.46
Al	3s3p ²	$^4P_{\frac{3}{2}} - ^4P_{\frac{1}{2}} = 75.78$	$^4P_{\frac{3}{2}} - ^4P_{\frac{1}{2}} = 46.58$	1.62
				1.67

† C. E. Moore, Atomic Energy Levels, *Natl. Bur. Standards Bull.* 467 (1949); W. E. Lamb, Jr., *Repts. Progr. in Phys.*, **14**, 19 (1951).

number of j values from $j = l + s$ to $j = l - s$ which is $2s + 1$ or $2l + 1$ depending on whether $l > s$ or $l < s$. Similarly the interval between successive j values for the same l and s is proportional to

$$(L \cdot S)_{j+1} - (L \cdot S)_j = \frac{\hbar^2}{2} [(j+1)(j+2) - j(j+1)] = \hbar^2(j+1)$$

This is known as the *Landé interval rule* for the separation of the components of a multiplet. Table 4.3 lists certain multiplet separations. The agreement of the preceding theory for one-electron atoms with experiment is very good. In the cases of the alkali doublets the separa-

tions are in general much greater than for a singly charged nucleus. While the separations cannot be calculated without a knowledge of the potential distribution caused by the electron structure, the general trends with n and Z are not greatly different from those predicted by the theory for a single electron. The Landé rule is seen to hold for the more complex multiplets listed.

4.3. One-electron Zeeman Effects. Certain effects of the presence of a magnetic induction upon atomic energy levels were seen in Sec. 4.1. There the discussion was limited to magnetic inductions that were either very small or very large in the sense that the change produced in the energy levels was much smaller or much larger than the fine structure of the levels of the atomic system concerned. If the magnetic induction was small, the coupling of the \mathbf{L} and \mathbf{S} vectors to form a resultant \mathbf{J} was but little perturbed and this angular-momentum vector precessed about \mathbf{B} as a result of the magnetic forces acting upon its associated magnetic moment. If the magnetic induction was large, the coupling between \mathbf{L} and \mathbf{S} was largely broken down and each precessed individually about \mathbf{B} with angular velocities characterized by the γ 's associated with the two types of angular momenta. A great deal of interest attaches to the intermediate or general case because it is the one most commonly realized experimentally and the precise quantitative information that it provides is invaluable in the determination of nuclear, atomic, and molecular properties.

The significant terms of the Schroedinger equation in the nonrelativistic approximation are of the three types that have been previously discussed. The first includes the kinetic- and potential-energy terms of the electron or electron structure, $V(r) + \mathbf{p}^2/2m_0$. The second type represents the energy of interaction between the magnetic moments associated with the spin and orbital angular-momentum vectors separately and the magnetic induction. These correspond to the terms in Eq. (4.9). The orbital term which comes from the expansion of $(\mathbf{p} - e\mathbf{A})^2$ is given explicitly by Eq. (2.21'). Writing the matrix \mathbf{S} for $(\hbar/2)\mathbf{s}$, these terms are seen to be $-(e/2m_0)(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B}$. The third type of term is that arising from the energy of interaction of the two types of angular momenta \mathbf{L} and \mathbf{S} as given by Eq. (4.10) or (4.11), $\eta\mathbf{S} \cdot \mathbf{L}$. The Hamiltonian equation for the energy levels of an electron structure with these angular momenta in a region of magnetic induction \mathbf{B} is thus

$$\mathcal{H}'\psi = \left[V(r) + \frac{\mathbf{p}^2}{2m_0} - \frac{e}{2m_0} \mathbf{L} \cdot \mathbf{B} - \frac{e}{m_0} \mathbf{S} \cdot \mathbf{B} + \eta \mathbf{S} \cdot \mathbf{L} \right] \psi = E'\psi \quad (4.12)$$

This is, of course, a matrix equation, and E' represents the matrix of the energy eigenvalues.

If $V(r)$ is a Coulomb potential, the function $\psi_{n,l,m}$ of Eq. (3.19) is known to be an eigenfunction of the equation consisting of the first two terms in Eq. (4.12) with the eigenvalues, or diagonal matrix elements, given by Eq. (3.20). Also, if \mathbf{B} is taken as being in the z direction for instance, the same eigenfunction is adequate if the first term in Eq. (4.12) containing \mathbf{L} is included, as can be seen from Eq. (3.18). The additional eigenvalues that are contributed by this term are $-(e\hbar B/2m_0)m_l$, where m_l takes the series of values from $+l$ to $-l$. Since the potential $V(r)$ is not of the Coulomb type if the atom contains more than one electron, the factor $R(r)$ of Eq. (3.19) is not satisfactory in general but the angular factors in ψ will remain the same if V is a function of r only. The interest here is in the energy-level differences arising from relative orientations of the vectors \mathbf{B} , \mathbf{L} , and \mathbf{S} , and it may be assumed that a satisfactory function of r in ψ can be determined, as say $R'(r)$, for the general case. Then $\psi = R'(r)\psi'$ is taken as a satisfactory solution of the equation consisting of the first three terms only of Eq. (4.12), and ψ' is a column matrix consisting of the tesseral harmonics $Y_{l,m_l}(\theta, \phi)$ for all values of m_l . The equation to be solved for the energy-level differences resulting from relative orientations of the three vectors in a spherically symmetrical potential and a magnetic induction \mathbf{B} can be written more simply in terms of ψ' as

$$H''\psi' = \left[\eta \mathbf{S} \cdot \mathbf{L} - \frac{e(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B}}{2m_0} \right] \psi' = E''\psi' \quad (4.12')$$

Here the eigenvalues E'' involve l , s , m_l , and m_s only, and the eigenfunction is written as ψ' again on the assumption, later seen to be justified, that a linear matrix function of the tesseral harmonics will be found to be a solution of Eq. (4.12').

The subsequent development becomes rather involved if large values of \mathbf{S} are permitted. \mathbf{S} will here be limited to the values characteristic of a single electron, that is, $\mathbf{S} = (\hbar/2)\boldsymbol{\sigma}$. This illustrates the general principles of the calculation and provides a number of results of immediate practical interest. Since \mathbf{S} is now assumed to be a matrix of two rows and two columns, a one-column two-row matrix to be operated on by \mathbf{S} must be associated with each Y_{l,m_l} in ψ' if the latter is to be a satisfactory eigenfunction for Eq. (4.12'). It is convenient to use the abbreviations $(+)$ and $(-)$ for the matrices $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ respectively, which

may be shown to be the spin portion of the eigenfunctions ψ' corresponding to $m_s = +\frac{1}{2}$ and $m_s = -\frac{1}{2}$, respectively. With \mathbf{B} in the direction of z , as was chosen earlier, the portion of the last term in

Eq. (4.12') involving \mathbf{S} is seen to contribute the eigenvalues ϵ and $-\epsilon$, where $\epsilon = -eB\hbar/2m_0$.

$$\frac{eB}{m_0} S_z(+) = \epsilon \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \epsilon \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \epsilon(+) \\ \frac{eB}{m_0} S_z(-) = \epsilon \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\epsilon \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\epsilon(-)$$

Thus $(+)$ and $(-)$ are satisfactory spin eigenfunctions with the eigenvalues corresponding to the energies of the two orientations of the spin $m_s = +\frac{1}{2}$ and $m_s = -\frac{1}{2}$ in the magnetic induction \mathbf{B} . Combining this result with that for the portion of the last term in Eq. (4.12') involving \mathbf{L} shows that the product of the column matrix Y_{l,m_l} times the matrix $(+)$ or the matrix $(-)$ is an eigenfunction of the equation consisting of the last term only of Eq. (4.12') with the energy eigenvalues

$$-\epsilon(m_l + 2m_s)$$

where m_l takes the integral values from $+l$ to $-l$ and m_s takes the values $+\frac{1}{2}$ and $-\frac{1}{2}$. The problem is then to see whether such an eigenfunction satisfies Eq. (4.12'), in which the first term is included as well.

This requires a brief development as follows:

$$\mathbf{S} \cdot \mathbf{L} = S_x L_x + S_y L_y + S_z L_z$$

and to avoid imaginary terms this will be written equivalently as

$$\mathbf{S} \cdot \mathbf{L} = S_z L_z + \frac{1}{2}[(S_x + iS_y)(L_x - iL_y) + (S_x - iS_y)(L_x + iL_y)]$$

It will be recalled from earlier equations in this section that the matrices for the x and y terms have entries only in the diagonals neighboring the main one. If the rows are labeled with the m_l 's, these matrices (regarded as transformations) connect a component m_l with the neighboring $m_l + 1$ or $m_l - 1$. Thus, for instance, using the elements for the angular-momentum matrices from Sec. 4.2,

$$L_x + iL_y = \begin{pmatrix} 0 & \hbar(2l)^{\frac{1}{2}} & 0 & 0 & 0 \\ 0 & 0 & \hbar[2(2l-1)]^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 0 & \hbar[3(2l-2)]^{\frac{1}{2}} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \\ L_x - iL_y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \hbar(2l)^{\frac{1}{2}} & 0 & 0 & 0 & 0 \\ 0 & \hbar[2(2l-1)]^{\frac{1}{2}} & 0 & 0 & 0 \\ 0 & 0 & \hbar[3(2l-2)]^{\frac{1}{2}} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

These relate the column matrices Y_{l,m_l} , the relation being between neighboring elements only. Thus

$$(L_x + iL_y)Y_{l,m_l} = \hbar[(l - m_l + 1)(l + m_l)]^{\frac{1}{2}}Y_{l,m_l-1}$$

$$(L_x - iL_y)Y_{l,m_l} = \hbar[(l - m_l)(l + m_l + 1)]^{\frac{1}{2}}Y_{l,m_l+1}$$

Since the S's are similar matrices with $s = \frac{1}{2}$ in place of l ,

$$(S_x + iS_y)(+) = \hbar(-), \quad (S_x + iS_y)(-) = 0$$

$$(S_x - iS_y)(+) = 0, \quad (S_x - iS_y)(-) = \hbar(+)$$

Finally, for the z components of the matrices,

$$L_z = \hbar \begin{pmatrix} l & 0 & 0 & 0 \\ 0 & (l-1) & 0 & 0 \\ 0 & 0 & (l-2) & 0 \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

or

$$L_z Y_{l,m_l} = \hbar m_l Y_{l,m_l}$$

and

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad S_z(+) = \frac{\hbar}{2}(+), \quad S_z(-) = -\frac{\hbar}{2}(-)$$

Thus the equations for $S \cdot L\psi'$ become

$$S \cdot L Y_{l,m_l}(+) = \frac{\hbar^2}{2} \{m_l Y_{l,m_l}(+) + [(l - m_l)(l + m_l + 1)]^{\frac{1}{2}} Y_{l,m_l+1}(-)\}$$

$$S \cdot L Y_{l,m_l}(-) = \frac{\hbar^2}{2} \{[(l - m_l + 1)(l + m_l)]^{\frac{1}{2}} Y_{l,m_l-1}(+) - m_l Y_{l,m_l}(-)\}$$

If m_j is written for the sum of m_l and m_s , the latter being $+\frac{1}{2}$ for $(+)$ and $-\frac{1}{2}$ for $(-)$, the equations become

$$S \cdot L Y_{l,m_j-\frac{1}{2}}(+) = \frac{\hbar^2}{2} m_l Y_{l,m_j-\frac{1}{2}}(+) + \frac{\hbar^2}{2} [(l - m_j + \frac{1}{2})(l + m_j + \frac{1}{2})]^{\frac{1}{2}} Y_{l,m_j+\frac{1}{2}}(-)$$

$$S \cdot L Y_{l,m_j+\frac{1}{2}}(-) = \frac{\hbar^2}{2} [(l - m_j + \frac{1}{2})(l + m_j + \frac{1}{2})]^{\frac{1}{2}} Y_{l,m_j-\frac{1}{2}}(+) - \frac{\hbar^2}{2} m_l Y_{l,m_j+\frac{1}{2}}(-)$$

m_j is the index of the component of the total angular momentum along B , and it is clear that both equations refer to the same value of m_j . With $m_l = m_j - \frac{1}{2}$ is associated $(+)$, or $m_s = \frac{1}{2}$, and with $m_l = m_j + \frac{1}{2}$ is associated $(-)$, or $m_s = -\frac{1}{2}$. The preceding equations are repre-

sented by the following simple two-row two-column matrix for any value of m_j :

$$\begin{pmatrix} \frac{\hbar^2}{2} (m_j - \frac{1}{2}) & \frac{\hbar^2}{2} [(l - m_j + \frac{1}{2})(l + m_j + \frac{1}{2})]^{\frac{1}{2}} \\ \frac{\hbar^2}{2} [(l - m_j + \frac{1}{2})(l + m_j + \frac{1}{2})]^{\frac{1}{2}} & -\frac{\hbar^2}{2} (m_j + \frac{1}{2}) \end{pmatrix} \quad (4.13)$$

In the particular case that $m_j = \pm(l + \frac{1}{2})$, the off-diagonal terms vanish.

Returning to Eq. (4.12'), the last term adds $\epsilon(m_l + 2m_s)$ or $\epsilon(m_j + m_s)$ along the main diagonal to the matrix of Eq. (4.13) times the factor η . Thus $H''\psi'$ for any m_j is given by

$$\begin{aligned} H'' \begin{pmatrix} \psi'_{m_j-\frac{1}{2}, \frac{1}{2}} \\ \psi'_{m_j+\frac{1}{2}, -\frac{1}{2}} \end{pmatrix} \\ = \begin{pmatrix} \eta \frac{\hbar^2}{2} (m_j - \frac{1}{2}) + \epsilon(m_j + \frac{1}{2}) & \eta \frac{\hbar^2}{2} [(l - m_j + \frac{1}{2})(l + m_j + \frac{1}{2})]^{\frac{1}{2}} \\ \eta \frac{\hbar^2}{2} [(l - m_j + \frac{1}{2})(l + m_j + \frac{1}{2})]^{\frac{1}{2}} & \eta \frac{\hbar^2}{2} (m_j + \frac{1}{2}) + \epsilon(m_j - \frac{1}{2}) \end{pmatrix} \\ \cdot \begin{pmatrix} \psi'_{m_j-\frac{1}{2}, \frac{1}{2}} \\ \psi'_{m_j+\frac{1}{2}, -\frac{1}{2}} \end{pmatrix} \end{aligned} \quad (4.14)$$

Here $\psi'_{m_j-\frac{1}{2}, \frac{1}{2}}$ is written for $Y_{l, m_j-\frac{1}{2}}(+)$ and $\psi'_{m_j+\frac{1}{2}, -\frac{1}{2}}$ for $Y_{l, m_j+\frac{1}{2}}(-)$.

Taking as a particular example $l = 1$, which corresponds to a p term,

$$\begin{array}{ccccccc} \psi' = Y_{1,1}(+) & Y_{1,0}(+) & Y_{1,1}(-) & Y_{1,-1}(+) & Y_{1,0}(-) & Y_{1,-1}(-) \\ m_j = \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{3}{2} \end{array}$$

If these functions are used and the symbols (+) and (-) are omitted for brevity, Eq. (4.14) with $\xi = \hbar^2\eta$ leads to the following six equations for Eq. (4.12'):

$$\begin{aligned} H'' Y_{1,1} &= \left(\frac{\xi}{2} + 2\epsilon \right) Y_{1,1} \\ H'' Y_{1,0} &= \epsilon Y_{1,0} + \frac{\xi}{\sqrt{2}} Y_{1,1} \\ H'' Y_{1,1} &= \frac{\xi}{\sqrt{2}} Y_{1,0} - \frac{\xi}{2} Y_{1,1} \\ H'' Y_{1,-1} &= -\frac{\xi}{2} Y_{1,-1} + \frac{\xi}{\sqrt{2}} Y_{1,0} \\ H'' Y_{1,0} &= \frac{\xi}{\sqrt{2}} Y_{1,-1} - \epsilon Y_{1,0} \\ H'' Y_{1,-1} &= \left(\frac{\xi}{2} - 2\epsilon \right) Y_{1,-1} \end{aligned} \quad (4.15)$$

The first and last of these equations are like those of Sec. 3.4 in that the Y 's appearing in them are satisfactory eigenfunctions and the energy eigenvalues are given by the parentheses on the right. Thus in these instances the eigenfunctions are satisfactory for the complete Eq. (4.12'). The second and third equations with $m_j = \frac{1}{2}$ and the fourth and fifth equations with $m_j = -\frac{1}{2}$ form two pairs, neither of which can be interpreted immediately in this way since these orbital states are mixed together by the matrix that formed the equations, and linear combina-

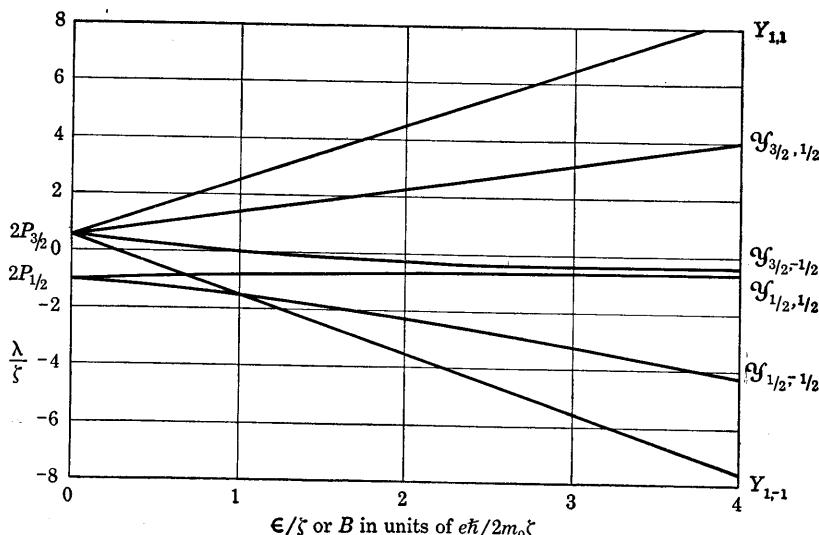


FIG. 4.5. Zeeman splitting of p -electron terms.

tions of the Y 's appearing in them must be sought as satisfactory eigenfunctions. As Eq. (4.12) is itself linear, such linear combinations will be satisfactory eigenfunctions of the equation as a whole. By means of a unitary transformation¹ formally similar to a simple rotation about an axis, satisfactory eigenfunctions may be formed in a way rather analogous to that in which an ellipse is referred to its principal axes by a rotation of coordinates. These linear combinations of the Y 's are written as script y 's; the identifying subscripts correspond to the j 's and the m_j 's involved. The former are identifiable a posteriori from the behavior of the levels when ϵ becomes small, as illustrated in Fig. 4.5. Thus for the first equation pair, which are equations two and three of Eq. (4.15),

$$\begin{aligned} y_{3/2} &= \cos \beta Y_{1,0} + \sin \beta Y_{1,1} \\ y_{1/2} &= \sin \beta Y_{1,0} - \cos \beta Y_{1,1} \end{aligned}$$

¹ L. I. Schiff, "Quantum Mechanics," p. 128, McGraw-Hill Book Company, Inc., New York, 1949.

Using this transformation, the angle β may be determined in such a way that the pair of equations for the Y 's becomes

$$\begin{aligned} H''Y_{\frac{1}{2}, \frac{1}{2}} &= \lambda_{\frac{1}{2}, \frac{1}{2}} Y_{\frac{1}{2}, \frac{1}{2}} \\ H''Y_{\frac{1}{2}, -\frac{1}{2}} &= \lambda_{\frac{1}{2}, -\frac{1}{2}} Y_{\frac{1}{2}, -\frac{1}{2}} \end{aligned}$$

where the λ 's are constants. Writing out the steps for the first of these equations, using the rotation above and its inverse for the Y 's in terms of the Y' 's,

$$\begin{aligned} H''Y_{\frac{1}{2}, \frac{1}{2}} &= H''Y_{1,0} \cos \beta + H''Y_{1,1} \sin \beta \\ &= (\epsilon Y_{\frac{1}{2}, \frac{1}{2}} \cos \beta - Y_{\frac{1}{2}, \frac{1}{2}} \sin \beta) \cos \beta \\ &\quad + \frac{\zeta}{\sqrt{2}} (Y_{\frac{1}{2}, \frac{1}{2}} \sin \beta + Y_{\frac{1}{2}, \frac{1}{2}} \cos \beta) \cos \beta \\ &\quad + \frac{\zeta}{\sqrt{2}} (Y_{\frac{1}{2}, \frac{1}{2}} \cos \beta - Y_{\frac{1}{2}, \frac{1}{2}} \sin \beta) \sin \beta \\ &\quad - \frac{\zeta}{2} (Y_{\frac{1}{2}, \frac{1}{2}} \sin \beta + Y_{\frac{1}{2}, \frac{1}{2}} \cos \beta) \sin \beta \\ &= \lambda_{\frac{1}{2}, \frac{1}{2}} Y_{\frac{1}{2}, \frac{1}{2}} \end{aligned}$$

β and $\lambda_{\frac{1}{2}, \frac{1}{2}}$ are determined from the fact that the sum of the terms containing $Y_{\frac{1}{2}, \frac{1}{2}}$ must vanish and the sum of the terms representing the coefficient of $Y_{\frac{1}{2}, \frac{1}{2}}$ must equal $\lambda_{\frac{1}{2}, \frac{1}{2}}$ by hypothesis. These conditions lead immediately to

$$\tan 2\beta = \frac{\sqrt{2}\zeta}{\epsilon + \zeta/2} \quad (4.16)$$

and

$$\begin{aligned} \lambda_{\frac{1}{2}, \frac{1}{2}} &= \epsilon \cos^2 \beta - \frac{\zeta}{2} \sin^2 \beta + \frac{\zeta}{\sqrt{2}} \sin 2\beta \\ &= \frac{1}{2} \left[\epsilon - \frac{\zeta}{2} + (\epsilon^2 + \epsilon\zeta + \frac{9}{4}\zeta^2)^{\frac{1}{2}} \right] \quad (4.17) \end{aligned}$$

Similarly the second equation of the pair leads to the same β , and

$$\lambda_{\frac{1}{2}, -\frac{1}{2}} = \frac{1}{2} \left[\epsilon - \frac{\zeta}{2} - (\epsilon^2 + \epsilon\zeta + \frac{9}{4}\zeta^2)^{\frac{1}{2}} \right] \quad (4.17')$$

Thus $Y_{\frac{1}{2}, \frac{1}{2}}$ and $Y_{\frac{1}{2}, -\frac{1}{2}}$ are eigenfunctions that are satisfactory for Eq. (4.12') if $m_j = \frac{1}{2}$, and the energy eigenvalues for these two functions are given by Eqs. (4.17) and (4.17'), respectively.

In an exactly similar way the fourth and fifth of Eqs. (4.15) can be transformed using

$$\begin{aligned} Y_{\frac{1}{2}, -\frac{3}{2}} &= -\cos \beta' Y_{1,-1} + \sin \beta' Y_{1,0} \\ Y_{\frac{1}{2}, -\frac{1}{2}} &= -\sin \beta' Y_{1,-1} + \cos \beta' Y_{1,0} \end{aligned}$$

and setting $H''Y_{\frac{1}{2},-\frac{1}{2}} = \lambda_{\frac{1}{2},-\frac{1}{2}}Y_{\frac{1}{2},-\frac{1}{2}}$ and $H''Y_{\frac{1}{2},\frac{1}{2}} = \lambda_{\frac{1}{2},\frac{1}{2}}Y_{\frac{1}{2},\frac{1}{2}}$. In the same way it is found that

$$\tan 2\beta' = \frac{\sqrt{2}\xi}{-\epsilon + \xi/2} \quad (4.16')$$

and the $\lambda_{-\frac{1}{2}}$'s may be derived from the corresponding $\lambda_{\frac{1}{2}}$'s by changing the sign of ϵ . Thus Eqs. (4.15) become

$$\begin{aligned} H''Y_{1,1} &= \left(\frac{\xi}{2} + 2\epsilon\right) Y_{1,1}, & H''Y_{\frac{1}{2},-\frac{1}{2}} &= \lambda_{\frac{1}{2},-\frac{1}{2}}Y_{\frac{1}{2},-\frac{1}{2}} \\ H''Y_{\frac{1}{2},\frac{1}{2}} &= \lambda_{\frac{1}{2},\frac{1}{2}}Y_{\frac{1}{2},\frac{1}{2}}, & H''Y_{\frac{1}{2},\frac{1}{2}} &= \lambda_{\frac{1}{2},\frac{1}{2}}Y_{\frac{1}{2},\frac{1}{2}} \\ H''Y_{\frac{1}{2},\frac{1}{2}} &= \lambda_{\frac{1}{2},\frac{1}{2}}Y_{\frac{1}{2},\frac{1}{2}}, & H''Y_{1,-1} &= \left(\frac{\xi}{2} - 2\epsilon\right) Y_{1,-1} \end{aligned} \quad (4.15')$$

These equations solve the problem completely for the particular case of a p electron for which $l = 1$. The generalization to the case in which the electron has a greater angular momentum, as for a d or f electron,

TABLE 4.4

Small-B energy-level shifts $\cos \beta = \cos \beta' = \sqrt{\frac{2}{3}}$ $\sin \beta = \sin \beta' = \sqrt{\frac{1}{3}}$	Eigenfunctions	Large-B energy-level shifts $\cos \beta = -\cos \beta' = 1$ $\sin \beta = \sin \beta' = 0$
$\frac{\xi}{2} + 2\epsilon$	$Y_{1,1} = Y_{\frac{1}{2},\frac{1}{2}}$	$\frac{\xi}{2} + 2\epsilon$
$\frac{\xi}{2} + \frac{2\epsilon}{3}$	$Y_{\frac{1}{2},\frac{1}{2}}$	ϵ
$-\xi + \frac{\epsilon}{3}$	$Y_{\frac{1}{2},\frac{1}{2}}$	$-\frac{\xi}{2}$
$-\xi - \frac{\epsilon}{3}$	$Y_{\frac{1}{2},-\frac{1}{2}}$	$-\epsilon$
$\frac{\xi}{2} - \frac{2\epsilon}{3}$	$Y_{\frac{1}{2},-\frac{1}{2}}$	$-\frac{\xi}{2}$
$\frac{\xi}{2} - 2\epsilon$	$Y_{1,-1} = Y_{\frac{1}{2},-\frac{1}{2}}$	$\frac{\xi}{2} - 2\epsilon$

follows immediately from Eq. (4.14). If the spin s exceeds $\frac{1}{2}$, the situation becomes considerably more complicated and a more general form of transformation must be resorted to for which the references should be consulted.

As an illustration of the use of Eqs. (4.15'), the results for a very small \mathbf{B} (initial Zeeman effect) may be written down by taking the first approximation in ϵ . Similarly the energies toward which the levels tend for very large \mathbf{B} (Paschen-Back effect) come from the values of the λ 's when ϵ exceeds ξ by a large amount. These limiting values are listed above and correspond to the nature of the curves which appear in Fig. 4.5

at the extreme left and extreme right. The values of β and β' give the Y eigenfunctions in terms of the Y 's. The level separations for large values of B differ from those given for the Paschen-Back effect in Eq. (4.3) only in the retention of the small term arising from the doublet separation $\frac{3}{2}\zeta$. The intermediate values of the energies associated with the different states of a p electron are plotted in Fig. 4.5 as a function of the ratio of the magnetic energy to the doublet separation energy. The behavior of the levels as B is varied, which is described by Eq. (4.15'), has been well verified experimentally in the optical region by Kent,¹ who made an extensive series of precise measurements on the Zeeman effect of the lithium doublet arising from the transition from the $2p$ to the $2s$ configuration. The molecular-beam spectroscopy in the microwave region, described in Sec. 3.3, provides a much more precise verification of the present theory, however, for this technique furnishes a means of directly measuring transitions between the sublevels of a given energy multiplet. The transitions that take place are of the magnetic-dipole-radiation type in which the magnetic-moment component along B may be thought of as changing without any change in the total orbital moment l . The selection rule is $\Delta m_j = 0, \pm 1$.

4.4. Hyperfine Structure. In the preceding discussion of one-electron atomic energy levels the nucleus of the atom has been considered as if it could be completely described as a point electric charge. This is a satisfactory first approximation, but a detailed study of atomic energy levels shows them to exhibit a complex structure several orders of magnitude smaller than that arising from the interaction between the orbital and spin moments of the electrons. This structure, known as *hyperfine structure*, is attributed to the influence of nuclear properties on the energy of the electron configuration. For many elements it is very complex and as yet incompletely resolved or understood. Certain components of it can be related to the finite mass and volume of the nucleus, and others to nuclear magnetic dipole and electric quadrupole moments. Contributions may also arise from a nonspherical shape and nonuniform distribution of mass and charge in the nucleus. For a discussion of these various points the references should be consulted; here the consideration will be limited primarily to the effect of nuclear magnetic moments.²

The study of hyperfine structure in optical spectra requires instruments of the highest possible resolving power and sources of radiation which minimize all causes of line broadening. Several ingenious arrangements have been devised. Jackson and Kuhn³ studied the hyperfine

¹ N. A. Kent, *Astrophys. J.*, **40**, 337 (1914).

² B. V. Rollin, *Repts. Progr. in Phys.*, **12**, 22 (1949); J. E. Mack, *Rev. Mod. Phys.*, **22**, 64 (1950); E. W. Foster, *Repts. Progr. in Phys.*, **14**, 288 (1951).

³ D. H. Jackson and H. Kuhn, *Proc. Roy. Soc. (London)*, **164A**, 48 (1938).

structure of sodium by an absorption technique. Sodium was heated at the lower ends of a vertical evacuated tube, and the vapor was condensed when it struck the top. Light from a sodium arc passed horizontally through this vertical atomic beam, and the broad emission lines were a background against which the absorption lines, which were very narrow because of the small velocity component in the line of sight, were observed. In their apparatus shown diagrammatically in Fig. 4.6, the Doppler broadening was less than 10^{-2} cm^{-1} and resolving powers as high as 5×10^6 could be used. Mundie and Meissner¹ studied the hyperfine structure of the emission spectrum of magnesium by a somewhat similar technique in which a narrow vertical beam of magnesium atoms from a furnace was excited by intense electron bombardment and the radiation emerging horizontally was analyzed by a Fabry-Perot interferometer. A wave-number precision of the order of 10^{-3} cm^{-1} was obtained. The contribution of nuclear moments to hyperfine structure can be studied with much greater precision in the microwave region, as described in Secs. 2.6 and 3.3. For a more complete description of this technique the references should be consulted.

An obvious cause of spectrum multiplicity which is generally classed as hyperfine structure is the presence of several isotopes of the element providing the source of the radiation. If more than one isotope is present, close lines appear of which certain groups can be identified with one isotope or another. Thus if hydrogen and deuterium are present, the difference in energy of terms of like total quantum number due to the difference in reduced mass is

$$\begin{aligned}\Delta E &= c\hbar R_\infty \left[\frac{m_r(\text{H})}{m_0} - \frac{m_r(\text{D})}{m_0} \right] \\ &\cong c\hbar R_\infty r_{ep} \left(1 - \frac{M_{\text{H}}}{M_{\text{D}}} \right)\end{aligned}$$

where r_{ep} is the electron-proton mass ratio ($1/1836.16$) and $M_{\text{H}}/M_{\text{D}}$ is the hydrogen-deuterium nuclear mass ratio ($\approx \frac{1}{2}$). Thus the fractional change in energy $\Delta E/E$ is of the order of 3×10^{-4} . Other changes in energy attributed to mass differences are observed in heavier polyelec-

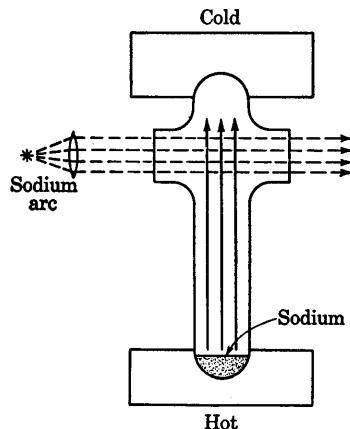


FIG. 4.6. Jackson and Kuhn absorption cell.

¹ L. G. Mundie and K. W. Meissner, *Phys. Rev.*, **65**, 265 (1944).

tronic atoms as well, but the calculation of the effect, which has been verified by observations on neon, is rather involved and will not be treated here.

A principal cause of hyperfine structure when but a single isotope is present is the energy of interaction between the electron structure and the higher-order moments such as an electric quadrupole moment or a magnetic dipole moment of the nucleus.¹ As has been seen earlier the principal interaction energy of a nuclear quadrupole moment would be expected to be with hydrogenic electron configurations of the *s* type for which $\psi\psi^*$ has a considerable value near the nucleus. These configurations are spherically symmetrical, however, so there is no dependence of the energy on the relative orientation of the nucleus and the quadrupole moment can rotate freely. For nonspherically symmetric electron configurations such as *p* and *d* types, nonuniform electric fields occur at the nucleus and if the nucleus has a quadrupole moment there is an orientation energy which, however, is generally only a few per cent of the magnetic dipole orientation energy.²

The interaction of the intrinsic nuclear magnetic moment and that of the external electron structure, which is responsible for the larger part of the orientation energy of the nucleus, can be calculated in much the same way as that of the spin-orbit term for an electron. The coupling term is analogous to that of Eq. (4.11), and, if \mathbf{J} is the electronic angular-momentum vector and \mathbf{I} the nuclear angular momentum where $\mathbf{I}^2 = \hbar^2 I(I + 1)$, $\eta' \mathbf{l} \cdot \mathbf{J}$ is the interaction energy. The constant η' may be shown to be $\eta g_i L^2 / Z J^2$ for a one-electron atom.³ Clearly the analogy is complete except for the coefficient η' . The number of hyperfine term components is $2I + 1$ or $2j + 1$, whichever is the smaller. Also the Landé interval rule applies, and the energy separation of neighboring hyperfine terms is $F + 1$, where $\mathbf{F}^2 = \hbar^2 F(F + 1)$ and $\mathbf{F} = \mathbf{J} + \mathbf{I}$ is the total angular momentum including that of the nucleus.

In a magnetic induction that would be very weak for the Zeeman effect the energy of interaction of the atomic moment and the magnetic induction is of the order of magnitude of that between \mathbf{J} and \mathbf{l} because of the smallness of the moment associated with \mathbf{l} . The magnetic interaction between \mathbf{B} and the \mathbf{J} and \mathbf{l} moments is $(g_j m_j + g_l m_l) \epsilon$, where $\epsilon = -e\hbar B / 2m_0$ from the earlier discussion. The coefficient g_l is the nuclear *g* value, which was seen in Sec. 2.6 to be a characteristic of a nucleus and smaller in order of magnitude than g_j by the ratio of the

¹ N. F. Ramsey, "Nuclear Moments," John Wiley & Sons, Inc., New York, 1953.

² G. E. Becker and P. Kusch, *Phys. Rev.*, **73**, 584 (1948); P. Kusch and H. M. Foley, *Phys. Rev.*, **74**, 250 (1948).

³ E. Fermi, *Z. Physik*, **60**, 320 (1930); G. Breit, *Phys. Rev.*, **37**, 51 (1931); S. A. Goudsmit, *Phys. Rev.*, **37**, 663 (1931).

electron to the proton mass. The quantity $m_i\hbar$ is of course the component of \mathbf{l} along \mathbf{B} . The effect of $\mathbf{l} \cdot \mathbf{J}$ coupling and a magnetic induction \mathbf{B} may be calculated just as was done for $\mathbf{l} \cdot \mathbf{S}$ coupling following Eq. (4.12'). Thus the pertinent terms are

$$H'_{IJ}\psi' = [\eta'\mathbf{J} \cdot \mathbf{l} + (g_j m_j + g_I m_I)\epsilon]\psi' \quad (4.18)$$

Just as s was taken as $\frac{1}{2}$ (one electron) in Eq. (4.13), j will for simplicity here be limited to $\frac{1}{2}$ corresponding to such terms as $^2S_{\frac{1}{2}}$, $^2P_{\frac{1}{2}}$, etc. Also, as $g_I \ll g_j$, g_I will be neglected in the following development. The m_j values are then $\pm \frac{1}{2}$ as for m_s in Eq. (4.13), I plays the role of l , and m_F represents m_j of Eq. (4.13). With these changes and adding the ϵ terms, the matrix analogous to Eq. (4.14) becomes

$$\begin{pmatrix} \frac{\zeta'}{2}(m_F - \frac{1}{2}) + \frac{g_j\epsilon}{2} & \frac{\zeta'}{2}[(I - m_F + \frac{1}{2})(I + m_F + \frac{1}{2})]^{\frac{1}{2}} \\ \frac{\zeta'}{2}[(I - m_F + \frac{1}{2})(I + m_F + \frac{1}{2})]^{\frac{1}{2}} & -\frac{\zeta'}{2}(m_F + \frac{1}{2}) - \frac{g_j\epsilon}{2} \end{pmatrix}$$

where ζ' is written for $\hbar^2\eta'$. As in Sec. 4.3 a linear combination of the angular eigenfunctions Y must be sought in terms of which the matrix is diagonal and the energy levels associated with the interactions are determined. This can be done as in Sec. 4.3 or more briefly by assuming the existence of such linear combinations as $y_{I \pm \frac{1}{2}, m_F}$ and writing the desired condition:

$$\begin{pmatrix} \frac{\zeta'}{2}(m_F - \frac{1}{2}) + g_j \frac{\epsilon}{2} & \frac{\zeta'}{2}[(I - m_F + \frac{1}{2})(I + m_F + \frac{1}{2})]^{\frac{1}{2}} \\ \frac{\zeta'}{2}[(I - m_F + \frac{1}{2})(I + m_F + \frac{1}{2})]^{\frac{1}{2}} & -\frac{\zeta'}{2}(m_F + \frac{1}{2}) + g_j \frac{\epsilon}{2} \end{pmatrix} \begin{pmatrix} y_{I+\frac{1}{2}, m_F} \\ y_{I-\frac{1}{2}, m_F} \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \cdot \begin{pmatrix} y_{I+\frac{1}{2}, m_F} \\ y_{I-\frac{1}{2}, m_F} \end{pmatrix}$$

If these homogeneous linear equations are written out, it is seen that the necessary condition that they shall have a solution is that the determinant of the matrix on the left with λ subtracted from the main-diagonal terms shall equal zero. This condition determines the possible values of λ as

$$\lambda_{F, m_F} = \frac{-\zeta'}{4} \pm \frac{1}{2}[\zeta'^2(I + \frac{1}{2})^2 + 2g_j m_F \zeta' \epsilon + g_j^2 \epsilon^2]^{\frac{1}{2}}$$

or writing $\Delta = \zeta'(I + \frac{1}{2})$, the doublet separation, and $\omega = g_j \epsilon / \Delta$, the ratio of the magnetic separation to the doublet separation, the diagonal energy terms are

$$\lambda_{F, m_F} = -\frac{\Delta}{2(2I + 1)} \pm \frac{\Delta}{2} \left(1 + \frac{4m_F}{2I + 1} \omega + \omega^2 \right)^{\frac{1}{2}} \quad (4.19)$$

This expression for the interaction energy of a nuclear magnetic moment with the moment of the surrounding electron structure in a $j = \frac{1}{2}$ state and in the presence of a magnetic induction \mathbf{B} was first derived by Breit and Rabi.¹ By analogy with the discussion of the atomic Zeeman effect, m_F takes all values from $+F = I \pm \frac{1}{2}$ to $-F = I \pm \frac{1}{2}$ in unit steps. The $(I + \frac{1}{2})$ series has two more terms in it than the $(I - \frac{1}{2})$ series, and the positive sign of Eq. (4.19) is associated with each of these. In analogy with the energy values associated with the eigenfunctions $Y_{1,1}$ and $Y_{1,-1}$ of Eq. (4.5), the energy values for these states are

$$\lambda_{F,\pm m_F} = -\frac{\Delta}{2(2I+1)} + \frac{\Delta}{2}(1 \pm \omega) \quad (4.19')$$

In the atomic-beam experiments transitions are induced between the energy states represented by the values of λ ; on the other hand, the hyperfine structure of spectral lines in the visible range is due to transitions

between these states and others of different electrostatic energy. As mentioned previously, the atomic-beam transitions correspond to magnetic dipole energy changes, and the radiation is of the magnetic rather than electric dipole type. The selection rules are quite similar to those for optical transitions, being

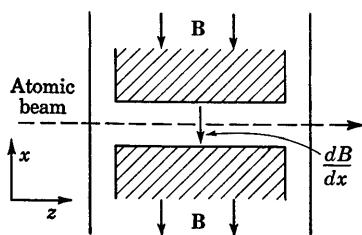


FIG. 4.7. Schematic representation of apparatus for zero-moment method.

$$\Delta F = 0, \pm 1 \text{ and } \Delta m_F = 0, \pm 1.$$

Equation (4.19) can be used with the atomic-beam technique to verify the correctness of associating an intrinsic angular momentum \mathbf{I} with a nucleus and also to determine the value of this angular momentum. If a beam of atoms in a $j = \frac{1}{2}$ state passes through a region of inhomogeneous magnetic induction, the force on an atom is given by the negative rate of change of the energy of the atom with respect to that coordinate, normal to the beam, along which \mathbf{B} varies. Thus if the beam is in the z direction and the gradient of the field is along x , as in Fig. 4.7, the force is given by

$$-\frac{d\lambda}{dx} = -\frac{d\lambda}{d\omega} \frac{d\omega}{dB} \frac{dB}{dx}$$

In general this results in a displacement of the atom along x , but it may be seen that for certain values of \mathbf{B} , $d\lambda/d\omega$ vanishes,

¹ G. Breit and I. I. Rabi, *Phys. Rev.*, **38**, 2082 (1931).

$$\frac{d\lambda}{d\omega} = 0 = \frac{\omega + \frac{2m_F}{2I+1}}{\left(1 + \frac{4m_F\omega}{2I+1} + \omega^2\right)^{\frac{1}{2}}} \quad \text{or} \quad \omega = -\frac{2m_F}{2I+1}$$

Thus the force in the x direction vanishes, and in such cases the atom is undeflected for those values of \mathbf{B} for which $\omega = -2m_F/(2I+1)$ [exclu-

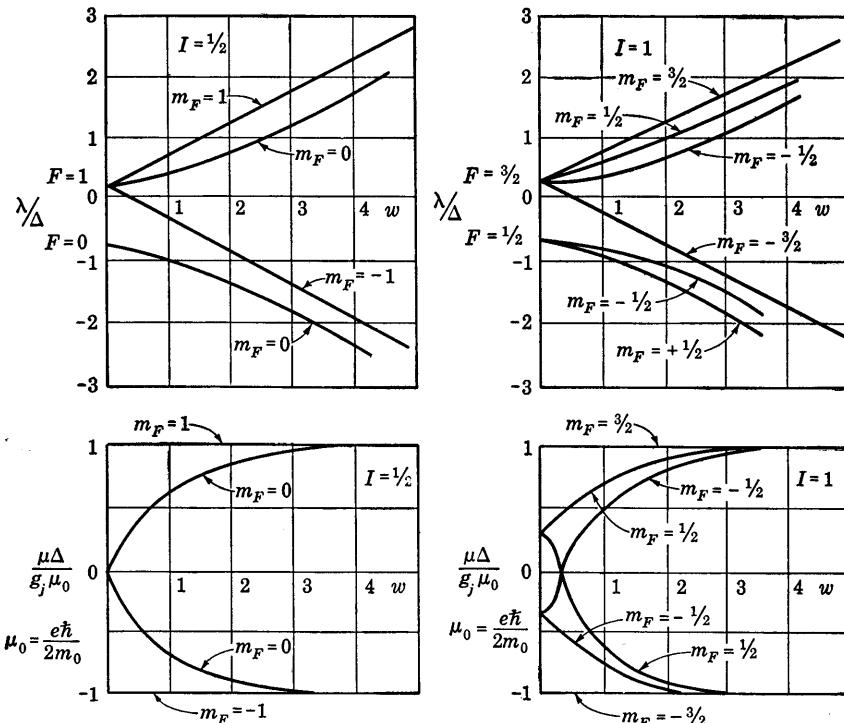


FIG. 4.8. Energy separations and magnetic moments as functions of ω or the magnetic induction.

sive of $m_F = \pm(I + \frac{1}{2})$ for then the denominator in the expression $d\lambda/d\omega$ vanishes as well]. The series of values of the magnitude of m_F are $I - \frac{1}{2}, I - \frac{3}{2}, I - \frac{5}{2}, \dots$, down to $\frac{1}{2}$ or 0 and the representative values of ω for which the force vanishes are given in Table 4.5 and illustrated in Fig. 4.8.

TABLE 4.5

$I = \frac{1}{2}$	$I = 1$	$I = \frac{3}{2}$	$I = 2$	$I = \frac{5}{2}$	$I = 3$	$I = \frac{7}{2}$
ω	0	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{3}{5}$	$\frac{2}{3}$	$\frac{5}{7}$

Thus if for say three values of B increasing from some very small value, maxima of transmission through the atomic beam apparatus are observed, the nuclear spin is either $\frac{5}{2}$ or 3; if one of these occurs at $B = 0$, the spin is $\frac{5}{2}$. This is known as the method of zero moments and was used in the middle 1930's by the Columbia group to determine the spin for the cesium nucleus ($I = \frac{5}{2}$) and for the potassium, sodium, and lithium nuclei ($I = \frac{3}{2}$). Molecular spectra provide an alternative technique for determining this important nuclear characteristic (Sec. 5.5). Magnetic-resonance techniques determine the magnetic moment-angular momentum ratio only, and hence a knowledge of I is essential to arrive at a value for the nuclear magnetic moments such as were given in Sec. 2.6.

TABLE 4.6

Atom	Atomic number	Isotope	Nuclear spin, \hbar	Nuclear moment $e\hbar/2M_p$		$g_I \times 10^4$
Li	3	6	1	+0.82189	± 0.00004	4.476
Li	3	7	$\frac{3}{2}$	+3.25586	± 0.00011	11.821
Na	11	22	3	+1.745	± 0.022	3.17
Na	11	23	$\frac{3}{2}$	+2.21711	± 0.00025	7.883
K	19	39	$\frac{3}{2}$	+0.391	± 0.001	1.42
K	19	40	4	-1.291	± 0.004	1.76
K	19	41	$\frac{3}{2}$	+0.215	± 0.001	0.78
Rb	37	85	$\frac{5}{2}$	+1.3532	± 0.0004	2.948
Rb	37	87	$\frac{3}{2}$	+2.7501	± 0.0005	9.985
Cs	55	133	$\frac{7}{2}$	+2.5771	± 0.0009	4.010
Cs	55	135	$\frac{7}{2}$	+2.7271	± 0.0033	4.244
Cs	55	137	$\frac{7}{2}$	+2.8397	± 0.0030	4.419

The molecular- and atomic-beam technique which was discussed briefly in Sec. 3.3 has contributed very extensively to our knowledge of nuclear magnetic moments as well as to our knowledge of nuclear angular momenta. When molecules are used in the beam the electron structure in general has no angular momentum or magnetic moment, that is, $j = 0$. In such a case Eq. (4.18) gives the energy values as $eg_I m_I$, and the phenomenon taking place may be thought of in terms of the simple precessional model for a moment which is invariant in magnitude as discussed in Sec. 2.6. Some of the most precise recent work has been done with atomic beams of the alkali metals and halogens for which $j = \frac{1}{2}$ and Eq. (4.19) applies. For the experimental details, which are of great interest, the references previously cited and those given below should be consulted.¹ To illustrate the results obtained, Table 4.6, taken from

¹ I. I. Rabi, S. Millman, P. Kusch, and J. R. Zacharias, *Phys. Rev.*, **55**, 526 (1939); D. R. Hamilton, *Am. J. Phys.*, **9**, 319 (1941); J. B. M. Kellogg and S. Millman, *Rev. Mod. Phys.*, **18**, 323 (1946); G. E. Pake, *Am. J. Phys.*, **18**, 438 (1950); N. F. Ramsey, "Nuclear Moments," John Wiley & Sons, Inc., New York, 1953.

Mack,¹ lists the present best values of the moments for the nuclei of the alkali atoms.

By the use of atomic-beam techniques, Prodell and Kusch² have made very precise measurements of the hyperfine-structure separations of the ground states of hydrogen and deuterium. These they find to be $\Delta\nu_H = 1420.4051$ and $\Delta\nu_D = 327.38424$ Mc, respectively, with an accuracy of about 0.2 part per million. Kusch and Foley³ using these techniques have shown that the magnetic-mechanical moment ratio γ_s for electron spin is $2(1 + \alpha/2\pi)\gamma_l$, where γ_l is the ratio for orbital motion [Eq. (2.23)]. Kusch and Foley made precise measurements of the ω 's of Eq. (4.19) by the atomic-beam method for certain states of known coupling of gallium, indium, and sodium. This involved measuring the radio frequencies which induce transitions between the states whose energy values are determined by the λ 's. The ratio of the ω 's for two states, say 1 and 2, yield the experimental ratios of the g 's for these states since $\omega = g_j\epsilon/\Delta$. From Sec. 4.1 $g_j = g_l(L/J) \cos LJ + g_s(S/J) \cos SJ$, where

$$\frac{L}{J} \cos LJ = \frac{j(j+1) + l(l+1) - s(s+1)}{2j(j+1)} = \alpha_l$$

$$\frac{S}{J} \cos SJ = \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)} = \alpha_s$$

If g_l is taken as being 1 and g_s is nearly 2, say $g_s = 2(1 + \delta_s)$ where δ_s is small,

$$\frac{\omega_1}{\omega_2} = \frac{g_{j1}}{g_{j2}} = \frac{\alpha_{l1} + 2(1 + \delta_s)\alpha_{s1}}{\alpha_{l2} + 2(1 + \delta_s)\alpha_{s2}} = \frac{2\alpha_{s1} + \alpha_{l1}}{2\alpha_{s2} + \alpha_{l2}} + \frac{\alpha_{s1}\alpha_{l2} - \alpha_{s2}\alpha_{l1}}{(2\alpha_{s2} + \alpha_{l2})^2} \delta_s$$

As the α 's are assumed known from the coupling, δ_s can be determined. This experiment yields a value in agreement with Eq. (2.23).

In addition to the modifications in the energy levels that are attributable to the interactions between the magnetic dipole and electric quadrupole moments of the nucleus with the electron structure, Lamb and Rutherford⁴ have recently observed a shift in the levels associated with an s electron that is due to a different cause. The experiments show that the 2S_1 level arising from the $2s$ electron configuration is displaced upward on a Grotrian diagram to smaller negative energies by an amount corresponding to 1062 ± 5 Mc, or about 3.54 wave numbers per meter from the 2P_1 level arising from the $2p$ electron configuration. In the

¹ J. E. Mack, *Rev. Mod. Phys.*, **22**, 64 (1950).

² A. G. Prodell and P. Kusch, *Phys. Rev.*, **88**, 184 (1952).

³ P. Kusch and H. M. Foley, *Phys. Rev.*, **74**, 250 (1948).

⁴ W. E. Lamb, Jr., and R. C. Rutherford, *Phys. Rev.*, **72**, 24 (1947); **81**, 252 (1951); W. E. Lamb, Jr., *Repts. Progr. in Phys.*, **14**, 19 (1951).

absence of hyperfine structure these two levels would be coincident as they correspond to the same total angular momentum ($j = \frac{1}{2}$). Figure 4.9 illustrates the situation approximately to scale. The experiments were performed by exciting the hydrogen atoms in an atomic beam by subjecting them to electron bombardment and noting that a small fraction of them had been excited to a state from which they could not return to the normal $1s$ configuration immediately by the emission of radiation. This was evidenced by the fact that these metastable atoms retained their energy of excitation until they impinged upon a tungsten plate

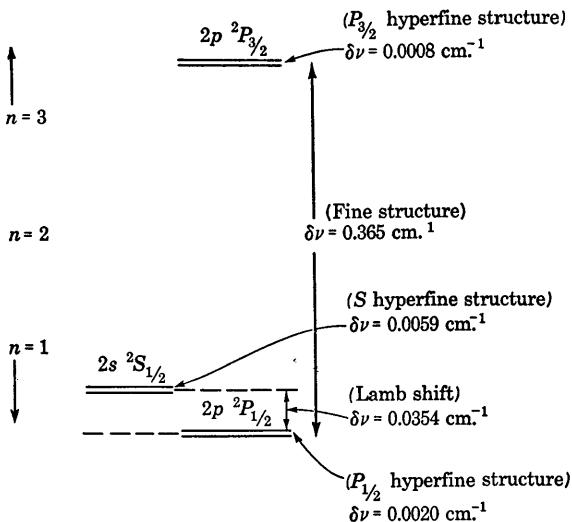


FIG. 4.9. Portion of Grotrian diagram for hydrogen levels, $n = 2$.

toward which the beam was directed. On impact with the plate these atoms gave up their energy of excitation to the free electrons in the metal which upon being drawn away to a positive electrode gave rise to an electric current that could be measured. The transition from the level $2s\ ^2S_{\frac{1}{2}}$ to the level $1s\ ^2S_{\frac{1}{2}}$, it will be recalled, is forbidden by the dipole selection rule as there is no change in the orbital angular momentum l . It was shown that this upper level was the metastable state involved, and, by producing a radio-frequency electromagnetic field near the beam, induced transitions from the $2s\ ^2S_{\frac{1}{2}}$ level to the $2p\ ^2P_{\frac{1}{2}}$ level were brought about as evidenced by the decrease in electron current from the tungsten plate. The resonant frequency at which the transitions are most effectively induced is a measure of the energy separation of the two levels.

The explanation of the raising of the s level is closely related to the interaction of an electron with an electromagnetic radiation field as

mentioned earlier in Secs. 1.8 and 2.6. The present theory of such fields is that they never fall completely to zero but that, in analogy with the so-called *zero-point oscillations* of a vibratory system such as a molecule (Sec. 5.4), there is a small but finite lower limit to the electromagnetic energy per unit volume which corresponds to the constant traversal of the volume by a continuous spectrum of radiation in which the amplitude of the electric vector is related to the frequency in the following way:

$$|E_\nu^2| = \frac{4\pi h}{\epsilon_0} \frac{\nu^3}{c^3} \quad (\text{see Sec. 6.5}).$$

As a result of this fluctuating electric field the electron in a hydrogen atom, for instance, tends to be randomly displaced from its strict equilibrium position centered on the nucleus. These oscillations of the electron reduce the time average of the energy with which the electron is bound to the nucleus, and the effect is most pronounced in the case of an *s* electron, the eigenfunction of which is concentrated around the nucleus as seen in Sec. 3.4. A calculation of the magnitude of the effect to be observed agrees to within about 1 per cent with the experimental observations, strongly supporting the correctness of the hypotheses involved.

The correctness of the theory of this fine-structure effect has been further confirmed by a study of the separation of the analogous levels in the case of singly ionized helium.¹ In this experiment ionization and excitation of the helium atoms in a low-pressure region were induced by a microsecond pulse of 200 ev electrons. After a brief interval to permit the decay of all but the metastable states, a pulse of microwave power was applied to induce the transition from the $2s\ ^2S_{\frac{1}{2}}$ metastable state of the $2p\ ^2P_{\frac{1}{2}}$ state. Return to the ground state $1s\ ^2S_{\frac{1}{2}}$ is allowed by a dipole radiation process, and the presence of the radiation associated with this transition was detected by liberation of photoelectrons by the 40.8-ev photons characteristic of the transition in ionized helium. The counters for detecting the photoelectrons were synchronized with the intervals during which the bursts of inducing radio-frequency power were applied. By this technique it was determined that the frequency of $14,021 \pm 60$ Mc was the most effective in inducing the $^2S_{\frac{1}{2}} \rightarrow ^2P_{\frac{1}{2}}$ transition and hence that the *P* state lies 5.798×10^{-5} ev above the *S* state. This numerical result is in excellent agreement with the fine-structure theory of the hydrogenic atom.

4.5. Identical Particles and the Pauli Principle. In the preceding sections the novel concept of spin, requiring formulation of the energy equation in terms of matrices for its utilization, has been introduced to provide a description of atomic phenomena. It is a fundamental concept as basic to atomic theory as the concepts of angular momentum and

¹ P. F. Yergin, W. E. Lamb, Jr., E. Lipworth, and R. Novick, *Phys. Rev.*, **90**, 377 (1953).

energy; unlike these latter it has not been taken over in a modified form from classical physics, but it represents a distinct and unique advance beyond any previous classical idea. In a somewhat analogous way the concept of identity may be endowed with a special meaning so that it connotes a much higher degree of equality between two entities than the word evokes in common usage. Two elementary particles of the same species, such as two electrons, are considered to be identical. By this is meant that it is in the nature of our physical world that two such particles have no intrinsic features that enable us to distinguish one from another and that it is meaningful to make statements about two such particles only if the quantum parameters or coordinates specifying their wave functions are not completely alike in the two cases. The coordinates or parameters need not be all different, but taken as two groups of specifications they cannot be identical in every respect. This is of course an extension of the customary use of the word identical which commonly applies to such like entities as twins, ball bearings, grains of sand, etc., which are clearly not identical in the sense of two electrons. In the first place, such objects cannot physically occupy the same space, though two electrons can, and, secondly, sufficiently minute examination will serve to identify each object and differentiate between two of them, but by hypothesis this is impossible for elementary particles such as electrons. In much the same way that the wave nature of matter precludes localization of elementary particles, identity (in the quantum-mechanical sense) prevents distinguishing between any two of them.

This extension of the concept of identity has very important implications for any quantum-mechanical description of a system containing two or more elementary particles of the same type. For instance, the function describing the system must be such that if the symbols associated with the two particles are interchanged there is no alteration in any observable quantity. This involves the existence of a high degree of symmetry for such a ψ function. A symmetry operation is by definition one which when applied to a system leaves it in a state indistinguishable from the initial state. It is a familiar concept in connection with crystal lattices which can be displaced by characteristic intervals, rotated, reflected, or inverted in certain classes of ways without any distinguishable alteration in the lattice as a result of such operations. The interchange of two electrons, say in an atom, is a symmetry operation in this sense. In consequence of this fact symmetry operations assume a particular importance in the description of atomic systems and phenomena.

The imagined interchange of two identical particles, i.e., a permutation, is a symmetry operation which leaves the system unchanged in any observable way. Consider the implication of this requirement on Eq. (3.6).

$$\mathsf{H}(q_1, q_2, \dots, q_n)\psi(q_1, q_2, \dots, q_n, t) = +i\hbar \frac{\partial}{\partial t} \psi(q_1, q_2, \dots, q_n, t) \quad (4.20)$$

where q_i represents all the coordinates and parameters necessary to specify the i th particle, and as implied above the system consists of n particles, some of which are identical. H , containing as it does squares of momenta and relative displacements and orientations of all the identical particles in the same way, is unaltered if the coordinates of two identical particles are interchanged. If ψ were changed by such a permutation, this could be only such a change as would leave the observable quantity $\psi\psi^* = |\psi|^2$ unaltered, i.e., if $P(1,2)$ represents the operation of hypothetically interchanging identical particles 1 and 2,

$$P(1,2)\psi(1,2, \dots, n) = \psi(2,1, \dots, n) = \pm\psi(1,2, \dots, n)$$

The choice of sign is a very significant matter since Eq. (4.20) shows that whichever sign applies at some arbitrary time will continue to apply thereafter. If the plus sign applies, ψ is said to be symmetric, and, if the minus sign applies, ψ is said to be antisymmetric. If ψ is symmetric, $\mathsf{H}\psi$ is also symmetric and therefore $(i\hbar \partial/\partial t)\psi$ is as well. If ψ is an analytic function of t , which it is assumed to be, the value of ψ at some time $t + \delta t$ is $\psi(t + \delta t) = \psi(t) + (\partial\psi/\partial t) \delta t$, where δt is assumed small. If $\psi(t)$ and $\partial\psi/\partial t$ are both symmetric, then $\psi(t + \delta t)$ is also and ψ will remain symmetric if it is so at some arbitrary time. Similarly, if ψ is antisymmetric at some time, it will be so at any later time. Thus symmetry or antisymmetry is a permanent property like charge or spin which can be attributed to an elementary particle. It is found that the elementary particles encountered in nature are of either one type or the other. Comparison of the predictions of the theory with experiment shows that a consistent description of atomic phenomena is given by the assignment of antisymmetric wave functions to all elementary particles which possess intrinsic spins that are odd multiples of $\frac{1}{2}\hbar$ and of symmetric wave functions to all particles which possess spins of zero or even multiples of $\frac{1}{2}\hbar$. Thus electrons, protons, neutrons, and possibly particles of small or no mass and no charge known as *neutrinos* and some particles of mass intermediate between electrons and protons known as *mesons* are to be described by antisymmetric wave functions; *photons* and other *mesons* are to be described by symmetric wave functions. This is one form in which the *Pauli principle* may be stated.

A satisfactory starting point for the polyelectronic problem is to assume as an approximation the same ψ functions that would be used to describe the states of the system in the absence of electronic interactions, namely, linear combinations of products of functions of the coordinates and parameters of each electron separately. A representative product

term in the sum would be $\psi_1(q_1)\psi_2(q_2)\psi_3(q_3) \dots$, in which an element in the product is a function involving the coordinates and parameters of one electron only. If by the symbol P is meant an arbitrary permutation of the q 's (coordinates and parameters) among the ψ 's for the different particles, the most general Ψ would be the $n!$ terms:

$$\Psi = \sum a_p P[\psi_1(q_1)\psi_2(q_2) \dots \psi_n(q_n)]$$

where the a_p 's are $n!$ arbitrary coefficients. If Ψ is to be symmetric all the a_p 's must be the same, whereas if Ψ is to be antisymmetric the a_p 's must be the same in magnitude but alternate in sign for successive permutations. Starting with the arbitrary assumption that the sign of the term with the factors in numerical order is plus, an odd number of pairwise interchanges for an antisymmetric function yields a term which is minus and an even number yields a term which is plus. From the definition of a determinant this is seen to be simply an alternative definition of an expanded determinant, so an antisymmetric wave function can be written as the determinant

$$\Psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(q_1) & \psi_2(q_1) & \psi_3(q_1) & \cdots & \psi_n(q_1) \\ \psi_1(q_2) & \psi_2(q_2) & \psi_3(q_2) & \cdots & \cdot \\ \psi_1(q_3) & \psi_2(q_3) & \psi_3(q_3) & \cdots & \cdot \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_1(q_n) & \psi_2(q_n) & \psi_3(q_n) & \cdots & \psi_n(q_n) \end{vmatrix} \quad (4.21)$$

where the factor $1/\sqrt{n!}$ is introduced to make $\int_v \Psi \Psi^* d\tau = 1$ if each of the individual ψ functions is so normalized that $\int_v \psi \psi^* d\tau = 1$. One property of a determinant is that the interchange of two rows or two columns changes the sign of the determinant, and this is in accordance with our definition of antisymmetry. Also, if two rows or two columns of a determinant are identical, the determinant vanishes. Thus the determinantal description of a Ψ function corresponds with the requirement that no two elementary particles of this type may be characterized by the same set of coordinates and quantum parameters. The above statement is also frequently said to be the Pauli principle; it is of course synonymous with the form previously given.

Periodic Table. An immediate major consequence of the foregoing discussion is the occurrence of periodic similarities in the sequence of chemical elements. It has been seen earlier that the electrons in an atom cannot be distinguished by their coordinates, so the only means of differentiating between them is by the assignment of quantum parameters determining their energies and angular momenta. It must be assumed that, because of interaction between the several electrons in an atom and between these electrons and external fields, the idealized situ-

ations, such as that encountered in the simplest description of hydrogen in Sec. 3.4 where the energy depended on but one parameter n , do not occur. This is physically realistic, and one would expect that it would be necessary to specify the angular momenta and their orientations for all the electrons in order to describe the atomic system completely since observable quantities such as the energy in general depend upon them. Since the spin is an intrinsic characteristic of an electron this means that the total quantum number n , orbital angular momentum l , and the angular-momentum projections m_l and m_s , for instance, must be given. The limitations resulting from the conclusion that these quantities cannot all be the same then determine permissible electron configurations. Assuming that the energy depends primarily on n and secondarily on l , the configurations of lowest energy for successive identical-electron addi-

TABLE 4.7

$n = 1$	$l = 0$	$m_l = 0$	$m_s = \frac{1}{2}$	1 electron $1s$
			$m_s = -\frac{1}{2}$	2 electrons $1s^2$
$n = 2$	$l = 0$	$m_l = 0$	$m_s = \frac{1}{2}$	3 electrons $1s^2 2s$
			$m_s = -\frac{1}{2}$	4 electrons $1s^2 2s^2$
	$l = 1$	$m_l = +1$	$m_s = \frac{1}{2}$	5 electrons $1s^2 2s^2 2p$
			$m_s = -\frac{1}{2}$	6 electrons $1s^2 2s^2 2p^2$
		$m_l = 0$	$m_s = \frac{1}{2}$	7 electrons $1s^2 2s^2 2p^3$
			$m_s = -\frac{1}{2}$	8 electrons $1s^2 2s^2 2p^4$
		$m_l = -1$	$m_s = \frac{1}{2}$	9 electrons $1s^2 2s^2 2p^5$
			$m_s = -\frac{1}{2}$	10 electrons $1s^2 2s^2 2p^6$
$n = 3$ etc.				

tions to a positive nucleus would be as shown in Table 4.7. The superscript indicates the number of electrons of that particular type, that is, $2s^2$ means two $2s$ electrons. In accordance with such a scheme no two electrons are specified by an identical combination of all four parameters, and hence the analysis is in conformity with the Pauli principle.

Though the formal analysis is consistent and exhaustive, the assignment of configurations to the particular ordering of m_l 's and m_s 's is not in general meaningful since it has here been chosen to assume that the energy of interaction with an external field predominates over internal interaction energies. However the interaction energies may be related, it is clear that there are at most two s electrons for $n = 1$, two s and six p electrons for $n = 2$, and two s , six p , and ten d electrons for $n = 3$, etc. These constitute the so-called *electron shells* of early chemical and X-ray nomenclature. Since the chemical behavior is determined principally by the least energetically bound electrons of the configuration it is clear that one would expect cyclic recurrences of chemical properties as one proceeds along the atomic sequence; such is indeed found to be the case. Table 4.8, shown on page 203, gives the lowest energy, or ground state, con-

figurations of the elements. It is seen that these are ordered much as indicated earlier till one reaches potassium where, instead of adding $3d$ electrons, $4s$ electrons represent the lower energy configuration. This is an effect of the dependence of energy on angular momentum in a non-Coulomb field (i.e., the Coulomb field of the nuclear charge screened by the atomic electrons) which becomes predominant at this stage; or alternatively stated, greater electrostatic energy is associated with the s type of electron concentrated near the nucleus of high atomic number than with the d type whose maximum charge concentration occurs at a finite radius. Thus the $3d$'s come in after calcium followed by $4p$'s with gallium, $5s$'s with rubidium, etc. The $4f$ electrons ($l = 3$) do not occur until on in the $5d$ group. This is the lanthanide or rare-earth group, all with very similar chemical properties since these properties are largely dependent on the barium- or lanthanumlike electronic structure common to them all. The electron configuration of an element can be written down by noting its atomic number and the electron positions occupied by the electrons of the elements of preceding numbers in the scheme. Thus selenium of atomic number 34 has the configuration of the preceding element for $n = 1, 2$, and 3 plus two $4s$ and four $4p$ electrons.

Selenium configuration $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^4$.

In the building of the d and f series, previously occupied positions are sometimes vacated. The configurations in these instances are indicated in parentheses. Thus in Pd^{46} both $5s$ electron positions are vacant.

Palladium configuration $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 4d^1$.

As will be seen in Sec. 5.3, association of atoms to form molecular combinations often occurs through sharing of pairs of electrons. This is called *covalent binding*. As the energy of such binding is usually only a few electron volts those electrons which are but loosely bound in the atomic structures are the ones able to participate in covalent bonds. Fig. 4.10 indicates the binding energy of the electrons of some of the lighter elements as determined from spectroscopic data. The atomic number is the abscissa and the energy on a logarithmic scale is the ordinate. It should be noted that these are inverted from the usual Grotrian-diagram convention. The continuous curves are not to be thought of as significant, but they carry the eye through the points characteristic of the discrete elements from 1 to 35. They indicate the way in which the loosely bound or optical electrons for low atomic numbers become the strongly bound or X-ray electrons for high atomic numbers. The energy differences between the curves are measures of the separation of the lowest terms represented by them. The K_{α} X-ray doublet for manganese has, for instance, a separation corresponding to the differences indicated by the arrows. At the other extreme the very close intersecting curves for $3d$ and $4s$ electrons between calcium and zinc indicate that the ground

TABLE 4.8. LOWEST ELECTRON CONFIGURATIONS OF THE ELEMENTS
OUTER-ELECTRON TYPES

<i>n</i>	<i>s</i>	<i>p</i>	<i>d</i>	<i>f</i>
1	1 H 2 He	5 B 6 C 7 N 8 O 9 F 10 Ne		
2	3 Li 4 Be	13 Al 14 Si 15 P 16 S 17 Cl 18 A	21 Sc 22 Ti 23 V ($4s^23d^3$) 24 Cr ($4s^13d^5$) 25 Mn ($4s^23d^5$) 26 Fe 27 Co 28 Ni ($4s^23d^8$) 29 Cu ($4s^13d^{10}$) 30 Zn ($4s^23d^{10}$) 39 Y 40 Zr ($5s^24d^2$) 41 Cb ($5s^14d^4$) 42 Mo ($5s^14d^5$) 43 Te ($5s^14d^6$) 44 Ru ($5s^14d^7$) 45 Rh ($5s^14d^8$) 46 Pd ($5s^04d^{10}$) 47 Ag ($5s^14d^{10}$) 48 Cd ($5s^24d^{10}$) 57 La 72 Hf 73 Ta 74 W 75 Re 76 Os ($6s^25d^6$) 77 Ir ($6s^05d^9$) 78 Pt ($6s^15d^9$) 79 Au ($6s^15d^{10}$) 80 Hg ($6s^25d^{10}$) 91 Pa ($6d^15f^2$) 92 U 93 Np 94 Pu 95 Am 96 Cm ($6d^15f^7$)	58 Ce ($5d^04f^2$) 59 Pr 60 Nd 61 Pm 62 Sm 63 Eu ($5d^04f^7$) 64 Gd ($5d^14f^7$) 65 Tb ($5d^04f^9$) 66 Dy 67 Ho 68 Er 69 Tm 70 Yb ($5d^04f^{14}$) 71 Lu ($5d^14f^{14}$)
3	11 Na 12 Mg			
4	19 K 20 Ca	31 Ga 32 Ge 33 As 34 Se 35 Br 36 Kr		
5	37 Rb 38 Sr	49 In 50 Sn 51 Sb 52 Te 53 I 54 Xe	81 T 82 Pb	
6	55 Cs 56 Ba	83 Bi 84 Po 85 At 86 Rn	89 Ac 90 Th	
7	87 Fr 88 Ra			

states for these elements are but little different in energy from the first excited configuration. When considered with the tabulation of the normal configurations these curves assist in accounting for the types of molecular species that the various atoms form. Two hydrogen atoms can pair their two valence electrons and form a molecule; two helium

atoms, on the other hand, do not. In helium the ionization potential is very high indicating that each electron is bound tightly to the atom in which the electrons are already paired so that the external pairing process does not take place. The ionization potential of beryllium is low, and

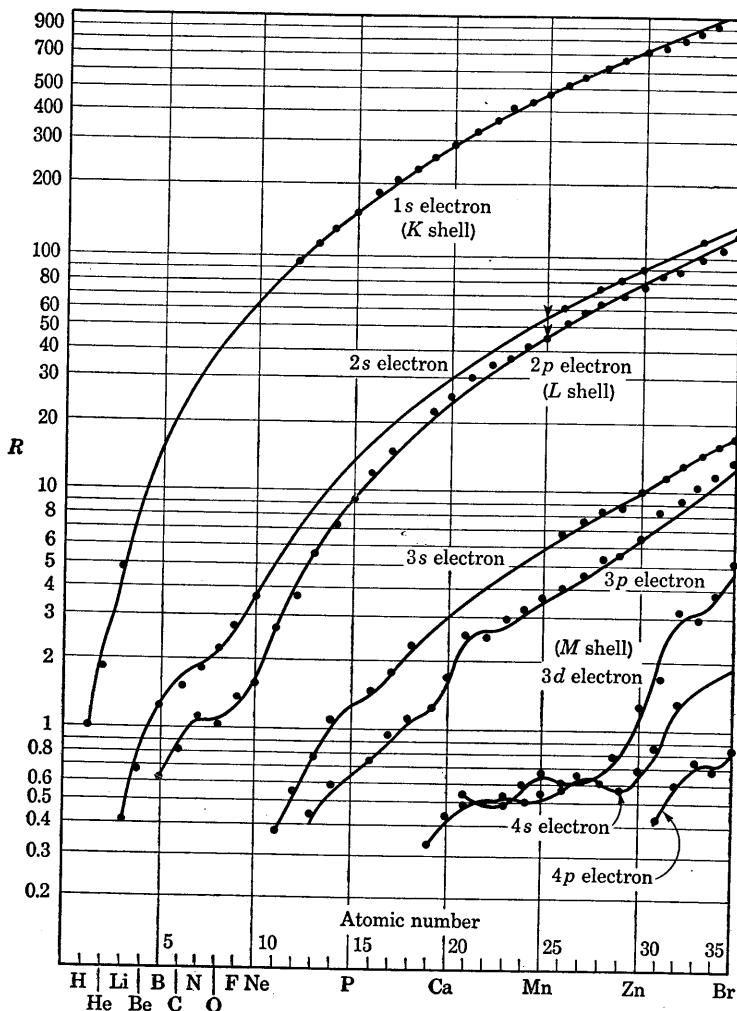


FIG. 4.10. Atomic electron energies. (Slater.) 1 rydberg = 13.53 ev = $1R$.

the configuration $1s^22s2p$ is but little higher in energy than the normal $1s^22s^2$. Thus in the former state, which is readily accessible, two electrons are available for pairing externally and a valence of two is observed. Similarly, though one would anticipate valences of 1 and 2 for boron and carbon, respectively, and CO with divalent carbon is observed, the

slightly excited configurations $1s^22s2p^2$ and $1s^22s2p^3$ for these two elements, respectively, are but little more energetic than the ground state, and in these configurations boron and carbon should and do display valences of 3 and 4, respectively. The case of carbon is particularly important because of the many compounds it forms, and it is interesting to note that only recently has indisputable spectroscopic evidence been found by Shenstone¹ for the value of the energy difference between the lowest states of the $2s^22p^2$ and the $2s2p^3$ configurations. This very important quantity in thermochemical theory is found to be 33,735.2 cm⁻¹, or 4.08 ev.

In this manner one may go through the table accounting qualitatively for chemical properties on the basis of electron configurations and the energies associated with them. The rare-gas configurations representing closed shells, He ($2s^2$), Ne ($2p^6$), A ($3p^6$), etc., followed by a large energy gap for the next electron, display great chemical inertness, as would be expected. Fluorine, chlorine, bromine, and iodine, on the other hand, represent these structures lacking one electron; two such atoms form a molecule essentially through the lowered energy represented by a structure in which a covalent electron pair is lacking. In the long series of elements in which d and f electrons are added, the terms arising from the different configurations overlap one another to a great extent and there is frequently very little energy difference between low terms. This is seen in the figure for the region between calcium and zinc. The analysis of these spectra has presented particular difficulties, and our knowledge of them is by no means complete. However, the various principles that have been educed in the course of this discussion appear adequate for formulating a description of all spectroscopic phenomena. The types of multiplets found are those that would be predicted from a knowledge of the angular momenta of the electrons making up the configuration. The electric dipole transitions that are observed to take place are those for which the orbital angular momentum changes by one unit. That is, a p electron becomes an s or a d electron at such transitions. The various regularities observed in isoelectronic sequences are all in general accord with expectation. The ordinate in Fig. 4.10 of the least strongly bound electron of an atom is the energy that would have to be supplied to it in order that it could leave the atom. This value is known as the *ionization potential*, and more precise values of this quantity for some of the lighter atoms are given in Table 4.9. It should also be mentioned that there is a small positive energy of binding of an additional electron to certain atoms. This quantity, known as the *electron affinity*, is important in many ionization and chemical processes in gases where the negative ion thus formed can play a significant role. For further information

¹ A. G. Shenstone, *Phys. Rev.*, **72**, 411 (1947).

on this very important but complex field the special references should be consulted.¹

TABLE 4.9

Ionization Potentials, ev (Moore)†

H	13.595	He	24.580
Li	5.390	Be	9.320
Na	5.138	B	8.296
		C	11.264
		N	14.54
		O	13.614
		F	17.42
		Ne	21.559
Mg	7.644	A	5.984
		Si	8.149
		P	11.0
		S	10.357
		Cl	13.01
		A	15.755

Electron Affinities, ev (Gaydon)‡

H	0.71	C	1.37	P	0.15	F	4.1
		N	0.04	S	2.06	Cl	3.75
		O	3.08			Br	3.6
						I	3.2

† C. E. Moore, *Atomic Energy Levels*, *Natl. Bur. Standards Bull.* 467 (1949).

‡ A. G. Gaydon, "Dissociation Energies," Chapman & Hall, Ltd., London, 1947.

Spectral Series. Aside from the multiplicity of spectrum lines the most arresting feature of atomic spectra is probably the occurrence of the recognizably related members of series, such as those of hydrogen that were mentioned in the preceding chapter. The alkali metals, the electron structure of which is characterized by one electron outside a closed *p* or *d* shell, exhibit similar series. Though they bear a close superficial resemblance to the hydrogen series, they differ in a number of important respects. These differences have their origin in the fact that the field in which the radiating electron exists is not a Coulomb field but is the field of the nucleus as modified by the spherically symmetric electron structure of the previously completed shells. The terms corresponding to large *l* values, which according to the discussion of Sec. 3.4 tend to represent the electron at the greatest average distance from the nucleus, resemble hydrogenic series most closely, for in these cases the field in which the electron generally finds itself is most closely akin to a Coulomb field from a central unit positive charge; the balance of the charge is effectively neutralized by the (*Z* - 1) electrons nearer the nucleus. On referring to Table 4.10 it can be seen that the term values for large *l* are closely the same as those for hydrogen for the same total quantum number *n*. As in the case of hydrogen the fine structure is not resolvable for these terms. Terms of lower *l* value represent configurations in which the electron concerned in the radiation process tends to spend a considerable portion of its time within the closed shells; in consequence, it is subject to an apparently greater nuclear charge and the binding energy increases. The binding energy is greatest for *l* = 0 for which electron

¹ H. E. White, "Introduction to Atomic Spectra," McGraw-Hill Book Company, Inc., New York, 1934; G. Herzberg, "Atomic Spectra and Atomic Structure," Prentice-Hall, Inc., New York, 1937; J. C. Slater, "Introduction to Chemical Physics," McGraw-Hill Book Company, Inc., New York, 1939; E. U. Condon and G. H. Shortley, "Theory of Atomic Spectra," Cambridge University Press, London, 1935.

TABLE 4.10. TERM VALUES FOR ALKALI-METAL CONFIGURATIONS, † cm⁻¹

$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$
Hydrogen $n(s,p,d,f)$						
109,679	27,419	12,186	6,854	4,387	3,046	2,238
Lithium $1s^2 n(s,p,d,f)$						
$l = 0$	43,486	16,280	8,475	5,188	3,500	2,535
$l = 1$	28,582	12,560	7,018	4,474	3,099	2,273
$l = 2$		12,203	6,863	4,390	3,047	2,237
$l = 3$			6,856	4,382	3,031	
Sodium $1s^2 2s^2 2p^6 n(s,p,d,f)$						
$l = 0$	41,449	15,709	8,248	5,074	3,437	
$l = 1$ ($j = \frac{1}{2}$ term)	24,493	11,182	6,409	4,153	2,909	
$l = 2$	12,276	6,900	4,412	3,062	2,249	
$l = 3$		6,860	4,390	3,040	2,231	
Potassium $1s^2 2s^2 2p^6 3s^2 3p^6 n(s,p,d,f)$						
$l = 0$	35,006	13,980	7,556	4,733		
$l = 1$ ($j = \frac{1}{2}$ term)	22,021	10,304	6,009	3,935		
$l = 2$ ($j = \frac{3}{2}$ term)	13,470	7,608	4,822	3,310		
$l = 3$	6,878	4,404	3,056	2,244		
Rubidium $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 n(s,p,d,f)$						
$l = 0$		33,689	13,558	7,378		
$l = 1$ ($j = \frac{1}{2}$ term)		21,110	9,974	5,854		
$l = 2$ ($j = \frac{3}{2}$ term)		14,334	7,989	5,002		
$l = 3$	6,898	4,418	3,068	2,252		
Cesium $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 5s^2 4d^{10} 5p^6 n(s,p,d,f)$						
$l = 0$			31,405	12,869		
$l = 1$ ($j = \frac{1}{2}$ term)			20,266	9,639		
$l = 2$ ($j = \frac{3}{2}$ term)		16,905	8,816	5,356		
$l = 3$ ($j = \frac{5}{2}$ term)	6,933	4,434	3,075	2,256		

† R. F. Bacher and S. Goudsmit, "Atomic Energy States," McGraw-Hill Book Company, Inc., New York, 1932.

type the probability of finding the electron in the immediate neighborhood of the nucleus is greatest. The succession of terms for a given l is again representable by the form $1/n'^2$, but n' is less than n by a characteristic constant, which is about 0.35 for a lithium s electron, 1.35 for a sodium s electron, 0.86 for a sodium p electron, etc., depending upon both Z and the angular momentum l . The doublet separations for low l values in the alkalis become very much greater than for hydrogen because of the direct dependence of energy upon angular momentum in a non-Coulomb field.

Quantitative calculations of atomic term values from electronic interactions are in general both lengthy and involved, and only the simplest examples illustrative of important general principles will be considered in a later section. However, certain general features can be immediately deduced from the foregoing. In the first place the possible total atomic angular momentum passes through periodic values as one proceeds through the s , p , d , and f series. At the conclusion of each such series the total angular momentum of the configuration is zero by the Pauli principle since all the possible orientations of these vectors are represented by electrons in the series. However, within the series a net angular momentum is possible and in general is present. Near the beginning of a series, i.e., near the tops of the divisions indicated by the inclined lines of Table 4.8, but few electrons are present of the particular type being added, and the dominant interaction between them is the scalar electrostatic potential energy due to their charges. The energy associated with relative orientations is very small. In the inevitable external magnetic induction due to neighboring atoms and electric currents, the orbital moments react and precess together as do the spin moments leading essentially to significant total l 's (L) and total s 's (S) for the atom as a whole. This is Russell-Saunders coupling previously mentioned, and on this assumption the particular terms designated by the symbol indicating the total S , L , and J arising from an electron configuration can be calculated.

This is very straightforward if there is no more than one electron of the same n and l concerned in the outer structure determining the term. Thus the excited beryllium configuration $2s2p$ has $S = \frac{1}{2} \pm \frac{1}{2}$, $L = 0 + 1$, and $J = 2, 1, 0$, and the terms arising from this configuration in accordance with the conventions of Sec. 4.1 are 1P_1 and $^3P_{2,1,0}$. If there are two or more electrons of the same type, that is, for which n and l are the same, the limitations imposed by the Pauli principle must be recognized in determining the possible relative orientations of their moments which in turn determine the resultant terms through L , S , and J . For instance, the moment orientations or components along a selected direc-

tion for the excited configuration of helium $1s2s$ are

m_{s1}	m_{s2}	m_{l1}	m_{l2}	S	L	Term
$\frac{1}{2}$	$\frac{1}{2}$	0	0	1	0	3S_1
$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0	1S_0

Whereas if the electrons have the same n (as in the normal state of helium), the configuration is $1s^2$ and the first of the above combinations is excluded since the two electrons would then be identically described in terms of n , l , m_l , and m_s and the only term arises from $n = 1$, $l = 0$, $m_{s1} = \frac{1}{2}$, $m_{s2} = -\frac{1}{2}$ or 1S_0 .

The determination of the number of different combinations of components and the associating of these together to form the possible terms are somewhat more lengthy as the number of electrons of a given type (n and l) and the magnitude of the angular momentum l increase. For tin, for instance, having two p electrons ($5p^2$) there are seen to be 15 possible combinations of m_s and m_l . These may be associated together in sequences S , $S - 1$, . . . , $-(S - 1)$, $-S$ and L , $(L - 1)$, . . . , $-(L - 1)$, $-L$ as indicated in Table 4.11 to yield the S and L values characterizing the resultant terms.

TABLE 4.11. TERMS FROM A p^2 CONFIGURATION

m_{s1}	m_{s2}	m_{l1}	m_{l2}	$m'_s = m_{s1} + m_{s2}$	$m'_l = m_{l1} + m_{l2}$	Class
$\frac{1}{2}$	$\frac{1}{2}$	1	0	1	1	a
$\frac{1}{2}$	$\frac{1}{2}$	1	-1	1	0	a
$\frac{1}{2}$	$\frac{1}{2}$	0	-1	1	-1	a
$\frac{1}{2}$	$-\frac{1}{2}$	1	1	0	2	b
$\frac{1}{2}$	$-\frac{1}{2}$	1	0	0	1	b
$\frac{1}{2}$	$-\frac{1}{2}$	1	-1	0	0	b
$\frac{1}{2}$	$-\frac{1}{2}$	0	1	0	1	a
$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0	c
$\frac{1}{2}$	$-\frac{1}{2}$	0	-1	0	-1	b
$\frac{1}{2}$	$-\frac{1}{2}$	-1	1	0	0	a
$\frac{1}{2}$	$-\frac{1}{2}$	-1	0	0	-1	a
$\frac{1}{2}$	$-\frac{1}{2}$	-1	-1	0	-2	b
$-\frac{1}{2}$	$-\frac{1}{2}$	1	0	-1	1	a
$-\frac{1}{2}$	$-\frac{1}{2}$	1	-1	-1	0	a
$-\frac{1}{2}$	$-\frac{1}{2}$	0	-1	-1	-1	a

In Table 4.11, class *a* includes all combinations for $L = 1$, $S = 1$, or $^3P_{2,1,0}$; class *b* for $L = 2$, $S = 0$, or 1D_2 ; class *c* for $L = 0$, $S = 0$, or 1S_0 . Therefore the terms arising from the configuration p^2 are 1S_0 , 1D_2 , and $^3P_{2,1,0}$. The extension of this tabular procedure to analyze such configurations as d^2 , p^3 , etc., for their resulting terms is obvious. The series

of terms for p^n and p^{6-n} are the same, as are also those for d^n and d^{10-n} , etc. Since L , S , and J are all zero for a completed shell, the ground term for a closed shell is 1S_0 .

In the regions of the electron configuration diagram representing nearly completed p , d , or f series, the spin-orbit interaction energy cannot be neglected in comparison with electrostatic interaction and L and S cease to be significant atomic parameters. This situation is known as jj coupling since the energies of interaction of the spin and orbital moments of each electron individually predominate over the mutual interactions of this type between two separate electrons. Individual electronic j 's remain significant as the components constituting the total atomic J which of course remains of importance. The nature of the configuration multiplets can be worked out in a straightforward way, but the Russell-Saunders nomenclature no longer applies.

4.6. Quantitative Methods; Helium. The general methods of quantitative approximation for obtaining the energy levels of polyelectronic atoms are partially empirical, not highly precise, and invariably laborious. For an account of them reference should be made to the specialized texts and original papers.¹ While the underlying theory appears to be entirely satisfactory, the calculations are so complex and the necessary approximations so numerous that ultimate precise quantitative agreement between calculated and observed energy levels has been achieved in relatively few cases. However, the principles of these approximation methods are very important and lead to the conception of nonclassical types of forces that play an essential role in atomic and molecular theory. Certain very simple instances will be considered in this section and the following chapter.

The property of *orthonormality*, which was described in Sec. 3.5, is an important characteristic of solutions of the wave equations. Consider the solutions u_n and u_m of the Schrödinger equation for two different allowed values of the energy E_n and E_m .

$$\frac{\hbar^2}{2m} \nabla^2 u_n - Vu_n = -E_n u_n$$

$$\frac{\hbar^2}{2m} \Delta^2 u_m^* - Vu_m^* = -E_m u_m^*$$

In Sec. 3.5 it was shown that

$$\int u_n u_m^* d\tau = \delta_{mn} \quad (4.22)$$

¹ L. I. Schiff, "Quantum Mechanics," chap. VII, McGraw-Hill Book Company, Inc., New York, 1949; W. V. Houston, "Principles of Quantum Mechanics," McGraw-Hill Book Company, Inc., New York, 1951; E. U. Condon and G. H. Shortley, "Theory of Atomic Spectra," Cambridge University Press, London, 1935.

where $\delta_{mn} = 0$ if $m \neq n$, $\delta_{mn} = 1$ if $m = n$. This condition of orthonormality is familiar for circular functions as well as for other harmonic functions. It enables one to expand certain functions in terms of the u 's just as functions are expanded in Fourier series in terms of sines and cosines.

Equation (4.22) is basic in the development of approximate methods of solving a Schroedinger equation for which an exact solution is not known. For instance, assume that a solution of a problem represented by the Hamiltonian H_0 is known. Such a solution is given by the u 's which satisfy the equation $H_0 u_n = E_n u_n$, where H_0 is the Hamiltonian for the hydrogen atom. Now if the problem is complicated by an additional term in the Hamiltonian, H' , so that the equation to be solved is

$$Hu = (H_0 + H')u = Wu$$

it is necessary to determine both the new permissible values of W as well as the functions u satisfying the equation with these values of W on the right. The method is to assume that H' is small and that to the first order the u 's depart but little from the u_n 's, and W but little from E_n . Thus

$$Hu \cong (H_0 + H')(u_n + u') = (E_n + W')(u_n + u')$$

where $u = u_n + u'$ and $W = E_n + W'$. If u' is considered to be expanded in a series of the u_n 's,

$$u' = \sum_j a_j u_j$$

Retaining only first orders in the quantities H' , u' , and W' , which are assumed small,

$$\begin{aligned} H_0 u_n + H_0 u' + H' u_n &= E_n u_n + E_n u' + W' u_n \\ H_0 \sum_j a_j u_j + H' u_n &= E_n \sum_j a_j u_j + W' u_n \end{aligned}$$

Multiplying this equation by u_k^* and integrating over the volume τ for which the u 's are orthonormal, the four terms yield the following:

$$\begin{aligned} \int u_k^* \sum_j a_j H_0 u_j d\tau &= a_k E_k \\ \int u_k^* E_n \sum_j a_j u_j d\tau &= a_k E_n \\ \int u_k^* W' u_n d\tau &= W' \delta_{nk} \\ \int u_k^* H' u_n d\tau &= H'_{kn} \quad (\text{definition}) \end{aligned}$$

Thus

$$\begin{aligned} a_k(E_n - E_k) + W' \delta_{nk} &= H'_{kn} \\ \text{or if } k = n, \quad W' &= H'_{nn} = \int u_n^* H' u_n d\tau \\ \text{if } k \neq n, \quad a_k &= \frac{1}{E_n - E_k} \int u_k^* H' u_n d\tau \end{aligned} \quad (4.23)$$

In this way the new permissible energy values $E_n + H'_{nn}$ and the coefficients for the u' series are obtained (except a_n which comes from normalizing u). If several u_n 's correspond to the same E , the above method

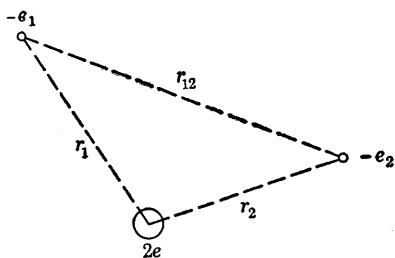


FIG. 4.11. Construction for the analysis of the two-electron problem.

a nucleus of charge $+2e$ and two electrons (see Fig. 4.11). As was seen earlier the major contributions to the potential energy come from the electrostatic interactions. These can be written

$$V = \left(-\frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{12}} \right) \frac{1}{4\pi\epsilon_0}$$

and the Schrödinger equation to this approximation is

$$\left[-\frac{\hbar^2}{2m} \nabla_1^2 - \frac{\hbar^2}{2m} \nabla_2^2 + \frac{1}{4\pi\epsilon_0} \left(-\frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{r_{12}} \right) \right] u = Wu \quad (4.24)$$

In the absence of the term $e^2/4\pi\epsilon_0 r_{12}$, a solution is known from Sec. 3.4 to be

$$u = u_{n,l,m}(1)u_{n',l',m'}(2)$$

where $u_{n,l,m} = Y_{l,m}R_n$, and (1) and (2) refer to electron number 1 and electron number 2, respectively. Thus to the zero approximation in the ground state ($n = 1$)

$$\begin{aligned} E_0 &= -2 \frac{m_0 Z^2 e^4}{2(4\pi\epsilon_0)^2 \hbar^2} = -8 \frac{m_0 e^4}{2(4\pi\epsilon_0)^2 \hbar^2} \\ &= -8R_\infty = -8 \cdot 13.6 = -108.8 \text{ ev} \end{aligned}$$

To be acceptable a wave function must be antisymmetric for the exchange of both coordinates and spin between the electrons. Though

would lead to infinite a_k 's in certain cases. This is called *degeneracy* and may be avoided by the choice of suitable linear combinations of the u_n 's as appropriate functions in terms of which to expand u' though in the present examples this necessity will not arise.

The foregoing technique may be used to solve approximately the problem of the helium atom consisting of

the spin does not occur in the Hamiltonian of Eq. (4.24) it is well to write the wave functions as if it did in order that the theory may apply more generally to cases where the energy is a function of spin orientation. In the notation of Sec. 4.3 spin orientations may be indicated by (+) and (-), where these are abbreviations for the matrices $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectively. The possibilities for two electrons as in helium are:

$$\begin{array}{llll} \text{State} & a & b & c & d \\ \text{Electron } 1 \uparrow = (+) = & \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \uparrow = (+) = & \begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \downarrow = (-) = & \begin{pmatrix} 0 \\ 1 \end{pmatrix}_1 \downarrow = (-) = \\ & \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 \downarrow = (-) = & \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 \uparrow = (+) = & \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \downarrow = (-) = \end{array}$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}_2$$

States *a* and *d* are symmetric for an interchange of electrons as they stand, and *b* and *c* can be combined to be either symmetric, *b* + *c*, or antisymmetric, *b* - *c*. The first symbol in parentheses is associated with electron 1 and the second symbol with electron 2:

$$b + c = (+-) + (-+) \quad b - c = (+-) - (-+)$$

Again as in Sec. 4.3 the contribution to the square of the magnitude of ψ comes from matrix multiplication of the transposed one-column matrices (i.e., one-row matrices, written under a wavy line) by the one-column matrices. Only terms with like subscripts, which refer to the same electron, operate on one another. For instance,

$$\begin{aligned} \widetilde{(b \pm c)(b \pm c)} &= [(10)_1(01)_2 \pm (01)_1(10)_2] \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}_2 \pm \begin{pmatrix} 0 \\ 1 \end{pmatrix}_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}_2 \right] \\ &= (1 \cdot 1 \pm 0 \cdot 0 \pm 0 \cdot 0 + 1 \cdot 1) = 2 \end{aligned}$$

For the functions describing the electrons, which are products of *u*'s and the (+) and (-) symbols, to be completely antisymmetric, symmetric *u*'s must be associated with *b* - *c* and antisymmetric *u*'s with *a*, *d*, and *b* + *c*. Thus the four following completely antisymmetric functions including spin are the ones that may be used in accordance with the Pauli principle.

Term type

$$\left. \begin{aligned} \frac{1}{\sqrt{2}} [u_n(1)u_m(2) + u_n(2)u_m(1)] \frac{1}{\sqrt{2}} [(+-) - (-+)] &\quad \text{Singlet} \\ \frac{1}{\sqrt{2}} [u_n(1)u_m(2) - u_n(2)u_m(1)] (++) \\ \frac{1}{\sqrt{2}} [u_n(1)u_m(2) - u_n(2)u_m(1)] \frac{1}{\sqrt{2}} [(+-) + (-+)] \\ \frac{1}{\sqrt{2}} [u_n(1)u_m(2) - u_n(2)u_m(1)] (--) \end{aligned} \right\} \text{Triplet} \quad (4.25)$$

These include numerical factors for normalization, as can be seen by forming the squares of the absolute magnitudes of the four functions.

The particular complicating energy term in Eq. (4.24), $H' = \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{12}}$, does not include a spin operator so one need not carry along the spin nomenclature. One need only recall that the negative sign in the combinations of u 's represents the three triplet states and the positive sign in this combination represents the singlet state.

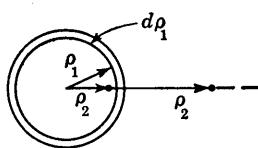


FIG. 4.12. Construction for the integration of the two-electron problem.

In the ground configuration, $1s^2$, for which both electrons are in the lowest state, corresponding to $n = 1$, the triplet wave functions vanish since the u 's cancel one another, and the normalizing factor for the spatial function is $\frac{1}{2}$.

$\frac{1}{2}[u_1(1)u_1(2) + u_1(1)u_1(2)] = u_1(1)u_1(2)$

$H' = e^2/4\pi\epsilon_0 r_{12}$ may be put with this in Eq. (4.23) to determine the first-order correction energy W' .

$$\begin{aligned} W' &= \frac{1}{4} \int_{\tau_1} \int_{\tau_2} [u_1(1)u_1(2) + u_1(2)u_1(1)]^2 \frac{e^2}{4\pi\epsilon_0 r_{12}} d\tau_1 d\tau_2 \\ &= \int_{\tau_1} \int_{\tau_2} [u_1(1)u_1(2)]^2 \frac{e^2}{4\pi\epsilon_0 r_{12}} d\tau_1 d\tau_2 \end{aligned}$$

From Sec. 3.4,

$$\begin{aligned} u_1 &= Y_{00}R_{10} = \frac{1}{\sqrt{4\pi}} 2 \left(\frac{2}{a_0}\right)^{\frac{3}{2}} e^{-2r/a_0} \\ W' &= \int_{\tau_1} \int_{\tau_2} \left[\frac{4}{4\pi} \left(\frac{2}{a_0}\right)^3 e^{-2(r_1+r_2)/a_0} \right]^2 \frac{e^2}{4\pi\epsilon_0 r_{12}} d\tau_1 d\tau_2 \end{aligned}$$

If $\rho_1 = 4r_1/a_0$, $\rho_2 = 4r_2/a_0$, $\rho_{12} = 4r_{12}/a_0$, and $d\tau = r^2 \sin \theta dr d\theta d\phi$,

$$\begin{aligned} W' &= \frac{64e^2}{4\epsilon_0\pi^3 a_0^6} \iint \frac{e^{-\rho_1} e^{-\rho_2}}{\rho_{12}} \frac{4}{a_0} \left(\frac{a_0}{4}\right)^6 \rho_1^2 \rho_2^2 d\rho_1 d\rho_2 \sin \theta_1 \sin \theta_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2 \\ W' &= \frac{e^2}{64\pi^3 \epsilon_0 a_0} \iint \frac{e^{-\rho_1} e^{-\rho_2}}{\rho_{12}} \rho_1^2 \rho_2^2 \sin \theta_1 \sin \theta_2 d\rho_1 d\rho_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2 \end{aligned}$$

The mutual repulsion of the two electrons increases ρ_{12} over what it would be if this were neglected. However, as a first approximation the displacements of the electron centers due to this cause will be neglected and it will be assumed that the charge distributions of each electron are symmetrically distributed about the nucleus with the density proportional to $e^{-\rho}$ as indicated. W' then represents the mutual electrostatic energy between these two symmetrical, interpenetrating charge distributions. The potential φ_1 due to a spherical shell of charge of radius ρ_1 and density

$(e/4\pi\epsilon_0)e^{-\rho_1}$ at a point 2 is, as indicated in Fig. 4.12,

$$\begin{aligned}\varphi_1 &= \frac{e}{4\pi\epsilon_0} \rho_1^2 e^{-\rho_1} \frac{1}{\rho_1} 4\pi d\rho_1 \quad (\rho_2 < \rho_1) \\ &= \frac{e}{4\pi\epsilon_0} \rho_1^2 e^{-\rho_1} \frac{1}{\rho_2} 4\pi d\rho_1 \quad (\rho_2 > \rho_1)\end{aligned}$$

Therefore the potential at the point 2 due to the entire distribution of charge 1 is

$$\begin{aligned}\varphi_1(\rho_2) &= \frac{e}{4\pi\epsilon_0} \frac{1}{\rho_2} 4\pi \left(\int_0^{\rho_2} \rho_1^2 e^{-\rho_1} d\rho_1 + \rho_2 \int_{\rho_2}^{\infty} \rho_1 e^{-\rho_1} d\rho_1 \right) \\ &= \frac{e}{4\pi\epsilon_0} \frac{1}{\rho_2} 4\pi \{ [-(\rho_1^2 + 2\rho_1 + 2)e^{-\rho_1}]_0^{\rho_2} + \rho_2[-(\rho_1 + 1)e^{-\rho_1}]_{\rho_2}^{\infty} \} \\ &= \frac{e}{4\pi\epsilon_0} \frac{4\pi}{\rho_2} [2 - (\rho_2^2 + 2\rho_2 + 2)e^{-\rho_2} + (\rho_2^2 + \rho_2)e^{-\rho_2}] \\ &= \frac{e}{4\pi\epsilon_0} \frac{4\pi}{\rho_2} [2 - (\rho_2 + 2)e^{-\rho_2}]\end{aligned}$$

This represents the integration over the coordinates indicated by the subscript 1. W' is then

$$\begin{aligned}W' &= \int \frac{\varphi_1(\rho_2) e 4\pi e^{-\rho_2} \rho_2^2 d\rho_2}{16\pi^2 a_0} \\ &= \frac{e^2}{4\pi\epsilon_0} \frac{4\pi e}{(4\pi)^2 a_0} \int_0^{\infty} \frac{4\pi}{\rho_2} [2 - (\rho_2 + 2)e^{-\rho_2}] e^{-\rho_2} \rho_2^2 d\rho_2 \\ &= \frac{e^2}{4\pi\epsilon_0} \frac{5}{4a_0}\end{aligned}$$

and as $a_0 = 4\pi\epsilon_0\hbar^2/m_0e^2$

$$W' = \frac{5}{4} \frac{m_0 e^4}{(4\pi\epsilon_0)^2 \hbar^2} = \frac{5}{2} R_{\infty} = 34.0 \text{ ev}$$

Thus, using the zeroth approximation earlier in this section, the energy of the ground state of helium is calculated to be

$$W = E_0 + W' = -108.8 + 34.0 = -74.8 \text{ ev}$$

The experimental value is 78.62 ev for the removal of both electrons from a helium atom. Since the energy of the helium ion is $-Z^2 R_{\infty}$ or $-4R_{\infty}$, the calculated ionization potential for one electron is

$$4R_{\infty} - \frac{5}{2}R_{\infty} = \frac{3}{2}13.6 = 20.4 \text{ ev}$$

as compared with the experimental value of 24.465 ev. More refined

calculations¹ using a variational method give a theoretical value of 24.461 ev, in excellent agreement with the observed value.

Though the illustrative calculation of the first approximation to the energy associated with the ground state of helium is of interest in itself, it is of less fundamental significance than the general form of the wave functions illustrated by Eqs. (4.25). The symmetry requirement imposed by the concept of identity, and the Pauli principle in the case of electrons, leads to an energy term (illustrated in next paragraph) that has no classical counterpart and is of essential importance in determining the energy of aggregation of elementary particles in atomic or molecular complexes. In general, the first approximation to the additional energy term arising from the joint presence of two electrons, which does not occur if each electron is present separately, is Eq. (4.23) with wave functions taken from Eqs. (4.25). This term represents a binding energy of these two electrons, and if one wishes to invoke the more intuitive concept of force its presence accounts for a force tending to determine the nature of the electron configuration. In the particular case first considered, $H' = e^2/4\pi\epsilon_0 r_{12}$ and $n = m = 1$ (electrostatic interaction only and $1s^2$ configuration), the spin portion of the wave function $[(+ -) - (- +)]$ was not affected by H' , and only the function symmetric in the coordinates was involved. In general, of course, H' depends on spin orientations (as for magnetic interactions) and n is not equal to m . Thus functions both symmetric and antisymmetric in the spatial coordinates must be used.

Consider briefly the same form of H' , that is, spin independent, but $n \neq m$ as for instance for the configurations $1s2s$, $1s2p$, etc. All wave functions of Eqs. (4.25) then have the normalization factor $1/\sqrt{2}$ as a result of summing over the possible spin orientations. Then

$$W' = \frac{1}{2} \iint [u_n(1)u_m(2) \pm u_m(1)u_n(2)] \frac{e^2}{4\pi\epsilon_0 r_{12}} [u_n^*(1)u_m^*(2) \pm u_m^*(1)u_n^*(2)] d\tau_1 d\tau_2$$

or $W' = J_{m,n} \pm K_{m,n}$, where

$$J_{m,n} = \iint \frac{e^2}{4\pi\epsilon_0 r_{12}} |u_n(1)u_m(2)|^2 d\tau_1 d\tau_2$$

$$K_{m,n} = \iint \frac{e^2}{4\pi\epsilon_0 r_{12}} u_n(1)u_m(2)u_m^*(1)u_n^*(2) d\tau_1 d\tau_2$$

The first of these is known as the *Coulomb integral* for obvious reasons. The quantity $J_{m,n}$ was calculated for the ground state of helium. The second term, $K_{m,n}$, which did not occur for the helium ground state as

¹ E. A. Hylleraas, *Z. Physik*, **65**, 209 (1930). See also H. A. S. Eriksson, *Nature*, **161**, 393 (1948).

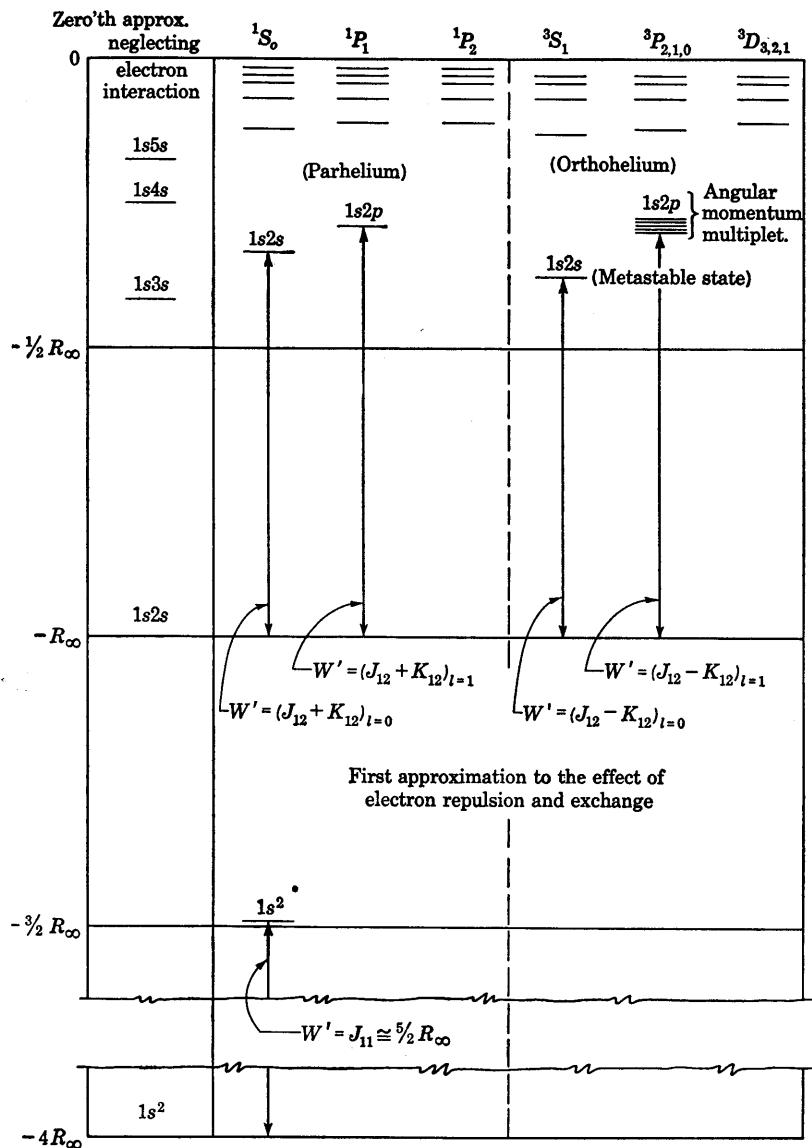


FIG. 4.13. Helium-energy-level scheme. (Paschen-Gotze.)

it is associated only with symmetric spins, is known as the *exchange integral*. It is evidently, in general, of the same order of magnitude as J and like it arises from the electrostatic-energy term $e^2/4\pi\epsilon_0 r_{12}$, but it has no classical analogue. $u_n(1)$ and $u^*(1)$ represent electron 1 first in the state specified by the quantum numbers n and secondly as simultaneously in a state represented by the quantum numbers m . A similar

statement can be made for $u_m(2)$ and $u_n^*(2)$ referring to the second electron. Thus this term, $K_{m,n}$, is traceable very directly to the identity of the two electrons which may be thought of as simultaneously or by frequent resonant interchange occupying both states n and m . The effect on the hydrogenlike terms of $W' = J \pm K$ is shown in Fig. 4.13. The zeroth-approximation hydrogenic states split into $S; S, P; S, P, D$; etc., states as n increases permitting increased l 's. The separations between the singlet and triplet energy levels are seen to be directly attributable to K . The fine structure of course comes from spin interactions not here considered.

PROBLEMS

4.1. Calculate the patterns for the initial Zeeman and extreme Paschen-Back effects for the spectrum lines resulting from the electron configuration changes $3d \rightarrow 2p$ and $1s2p \rightarrow 1s^2$. Assume that the selection rules are strictly observed.

4.2. The first members of the so-called principal-series lines in the alkali-metal spectra arise from an electron configuration change $p \rightarrow s$. The wave-number separation of the double level associated with the p electron configuration is 17.2 cm^{-1} for Na, 57.9 cm^{-1} for K, and 237.7 cm^{-1} for Rb. Assuming that the initial Zeeman pattern is observed for a magnetic induction such that the total extent of the spectrum pattern is less than 5 per cent of the doublet separation, calculate the maximum value of B which can be used in the three instances to observe this pattern. Assuming that the extreme Paschen-Back effect is observed for values of B such that the total extent of the spectrum pattern is greater than the doublet separation, calculate the minimum value of B that would be necessary to observe this pattern in the three instances.

4.3. Drawing upon the polarization discussion of Secs. 3.4 and 4.1, describe the pattern that would be observed on studying the initial Zeeman and Paschen-Back effects by viewing the light emitted from a $p \rightarrow s$ transition (a) parallel to the direction of B , (b) normal to the direction of B . If in addition to direct observation a nicol prism is used to study the light emitted normal to B and a quarter-wave plate to study that emitted parallel to B , what would be the nature of the polarizations found for the several lines?

4.4. Verify the matrix equations

$$\begin{aligned} \sigma_x\sigma_x - \sigma_x\sigma_x &= a_y\sigma_y - \sigma_ya_y = a_z\sigma_z - \sigma_za_z = 0 \\ \sigma_x\beta - \beta\sigma_x &= \sigma_y\beta - \beta\sigma_y = \sigma_z\beta - \beta\sigma_z = 0 \\ \sigma_xa_y - a_y\sigma_x &= 2ia_z \\ \sigma_ya_z - a_z\sigma_y &= 2ia_x \\ \sigma_za_x - a_x\sigma_z &= 2ia_y \end{aligned}$$

4.5. Form the scalar product of an arbitrary vector r and the matrix δ , and show that the absolute magnitude of this product is $|r|$. Show also that $r^2\sigma^2 = 3(r \cdot \delta)^2$.

4.6. Show that the transformation matrix representing a 90° rotation in a positive sense about the z axis is $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Show by matrix multiplication that two successive 90° rotations are equivalent to one 180° rotation, and that three successive 180° rotations, one about each axis, are equivalent to no rotation. A reflexion in a coordinate plane is a transformation in which one axis is inverted in sign, i.e., x goes into $-x$

for a reflexion in the yz plane. Show by the use of matrices that an even number of reflexions is equivalent to a rotation and that an odd number of reflexions cannot be so represented.

4.7. Three resistances are joined in the form of a triangle, one apex of which is grounded. If the potential differences from ground of V_1 and V_2 are applied to the other two apexes, show that the matrix representing the network which when multiplied by $\begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ yields the current to the apexes $\begin{pmatrix} i_1 \\ i_2 \end{pmatrix}$ is

$$\begin{pmatrix} S_{10} + S_{12} & -S_{12} \\ -S_{12} & S_{20} + S_{12} \end{pmatrix}$$

where the S 's are the conductances of the branches and the subscripts designate the junctions terminating them.

4.8. Extending Prob. 4.7, consider an impedance network in the form of two triangles having one side in common with a common apex grounded. Write the matrix representing the network, i.e., relating the current flowing to the three other apexes, and the potential of each, and show that this is fundamentally different from the L matrices for $l = 1$ in the sense that for positive values of the S 's it cannot be represented by $aL_x + bL_y + cL_z$ for any choice of a , b , and c .

4.9. Show that the L matrices for any value of the total angular-momentum index l discussed in Sec. 4.2 satisfy the angular-momentum equations (3.17).

4.10. A cartesian coordinate system x , y , z is rotated through an angle ψ about a line having direction cosines a , b , and c with the initial axes. The final positions of the axes, or the new rotated coordinate system, is x' , y' , z' . Show that the matrix of transformation from the unprimed to the primed axes is

$$\begin{pmatrix} a^2(1 - \cos \psi) + \cos \psi & ba(1 - \cos \psi) - c \sin \psi & ca(1 - \cos \psi) - b \sin \psi \\ ab(1 - \cos \psi) - c \sin \psi & b^2(1 - \cos \psi) + \cos \psi & cb(1 - \cos \psi) + a \sin \psi \\ ab(1 - \cos \psi) + b \sin \psi & bc(1 - \cos \psi) - a \sin \psi & c^2(1 - \cos \psi) + \cos \psi \end{pmatrix}$$

What is the matrix if two of the direction cosines vanish? What is the value for an infinitesimal rotation?

4.11. Write out Eq. (4.7) as four homogeneous linear equations among the function components ψ_1 , ψ_2 , ψ_3 , and ψ_4 for a free particle ($V = 0$) and show that the vanishing of the determinant of the coefficients, which is the condition that a solution exists, implies that $E^2 - c^2 p^2 - m_0^2 c^4 = 0$.

4.12. Show that the sum of the frequencies of the ten spectrum lines observed in accordance with the selection rules ($\Delta j = \pm 1, 0$; $\Delta m_j = \pm 1, 0$) for a transition in a single-electron spectrum from the $2p$ configuration to the $1s$ configuration in the presence of an arbitrary value of magnetic induction is less than ten times the frequency corresponding to the electrostatic energy separation of the configurations by an amount equal to the doublet energy interval of the $2p$ state.

4.13. Starting from Eq. (4.14), show that the Zeeman levels for a one-electron spectrum are given by

$$\lambda_{j,m_j} = \frac{\Delta}{2} \left[-\frac{1}{2l+1} + 2m_j w \pm \left(1 + \frac{4m_j w}{2l+1} + w^2 \right)^{\frac{1}{2}} \right]$$

where $\Delta = \xi(l + \frac{1}{2})$, $w = \epsilon/\Delta$, and m_j , l , ϵ , and ξ have the same significance as in the text. The most direct way to do this is to assume that eigenfunctions can be found by a transformation which will have the matrix eigenvalues λ and to determine the permissible values of λ for the condition that the determinant of the coefficients must vanish for a solution to exist.

4.14. Using the results of the preceding problem, show that the sum of all the energy-level shifts brought about in the Zeeman splitting of a one-electron configuration is zero for any value of the magnetic induction.

4.15. Draw the curves for the variation of the hyperfine-structure energy levels in units of Δ for an atom with $j = \frac{1}{2}$ and $I = \frac{3}{2}$ as a function of w . Draw the curves for the effective magnetic moment in units of $g_J\mu_0/\Delta$ for such an atom as a function of w .

4.16. Using Eq. (4.19), show that a group of transitions between the hyperfine-structure levels all approach the value $\Delta/(2I + 1)$ for large values of w . Show that the sum of the energy differences for this particular series of transitions is equal to Δ for any value of w .

4.17. Show that if the relatively small nuclear g value is not neglected, as was done in arriving at Eq. (4.19), this equation becomes

$$\lambda_{F,m_F} = -\frac{\Delta}{2(2I+1)} \pm \frac{\Delta}{2} \left(1 + \frac{4m_F}{2I+1} w + w^2\right)^{\frac{1}{2}} + \epsilon g_I m_F$$

where $\Delta = \zeta'(I + \frac{1}{2})$, $w = (g_J - g_I)\epsilon/\Delta$, and I , m_F , ϵ , and ζ' have the same significance as in the text.

4.18. In the theory of magnetic materials it is shown that the magnetic induction in a spherical cavity hollowed out in a medium of magnetic moment per unit volume μ_v is $-2\mu_v/3\epsilon_0 c^2$. Using this result and assuming that the nucleus of a hydrogenic atom in an s state occupies such a spherical cavity in the electron distribution, show that the energy separation between the hyperfine-structure levels for an s state of quantum number n is given by

$$\delta U_{I,S} = \frac{4}{3} hc R_\infty \alpha^2 \left(\frac{Z}{n}\right)^3 g_I (2I + 1)$$

I is the nuclear-spin quantum number and g_I is the nuclear g value in units of $e\hbar/2m_0$. Using the result of Prob. 3.14, show that the ratio of the quadrupole energy-level shift for an s electron to the hyperfine-structure separation is $(2\pi^2/\lambda_c^2) QZ/[(2I + 1)g_I]$.

4.19. Show by averaging the third term of Eq. (4.10) over the atomic volume that it yields an s -term fine-structure energy of $\delta U_S = hc R_\infty \alpha^2 (Z^4/n^3)$ for a hydrogenic atom of nuclear charge Ze in a state of total quantum number n .

4.20. Show that a classical calculation of the magnetic energy of an electron-spin moment $\mathbf{u}_s = (e/m_0)\mathbf{S}$ at a distance r from a nucleus of positive charge Ze moving with a velocity \mathbf{v} , or angular momentum $M\mathbf{r} \times \mathbf{v}$, yields

$$U_m = \frac{1}{4\pi\epsilon_0 c^2} \frac{1}{r^3} \frac{Ze}{M} \frac{e}{m_0} \mathbf{L}_M \cdot \mathbf{S}$$

or in terms of the negatively charged electron's orbital angular momentum \mathbf{L} ($\mathbf{L}_M/M = \mathbf{L}/m_0$)

$$U_m = -\frac{1}{4\pi\epsilon_0 c^2} \frac{1}{r^3} Z \left(\frac{e}{m_0}\right)^2 \mathbf{L} \cdot \mathbf{S}$$

On comparing this with the last term of Eq. (4.10), it is evident that such a classical calculation is in error by the factor $-\frac{1}{2}$. If it is assumed as a result that a correct value will be obtained for the interaction energy between atomic vectors if \mathbf{u}_s is taken as $-\frac{1}{2}(e/m_0)\mathbf{S}$, show that the mean magnetic induction that would be produced at the nucleus because of the magnetic moments associated with the \mathbf{L} and \mathbf{S} vectors of a single electron which precess about their common resultant \mathbf{J} would be

$$\mathbf{J} \frac{\mathbf{J} \cdot (\mathbf{B}_{\mu_L} + \mathbf{B}_{\mu_S})}{J^2} = \frac{1}{4\pi\epsilon_0 c^2} \frac{e}{m_0} \frac{1}{r^3} \frac{L^2}{J^2} \mathbf{J}$$

where the result of Prob. 4.5 has been used. Show that, in consequence, the hyperfine-structure energy for a nucleus of magnetic moment $g_I(e/2m_0)\mathbf{I}$ within such an electron configuration is

$$U_{I,J} = \eta' \mathbf{I} \cdot \mathbf{J}$$

where

$$\eta' = \frac{2\pi R_\infty c \alpha^3 g_I Z^3}{\hbar j(j+1)(l+\frac{1}{2})n^3}$$

4.21. If in calculating η the difference between the reduced mass of the atomic system and m_0 is not neglected, show that the hyperfine-structure wave-number separation is given by

$$\delta\tilde{\nu} = \frac{R_\infty \alpha^2 g_I Z^3}{(l+\frac{1}{2})j(j+1)n^3} \left(\frac{m_r}{m_0}\right)^3 (I + \frac{1}{2})$$

Show that the ratio of the ground-state separations for hydrogen and deuterium is given by

$$\frac{\delta\tilde{\nu}_H}{\delta\tilde{\nu}_D} = \frac{4}{3} \frac{\mu_D}{\mu_H} \left(\frac{m_{rH}}{m_{rD}}\right)^3$$

The values of the ratios [$(\delta\tilde{\nu}_H/\delta\tilde{\nu}_D) = 4.3386484$ and $(\mu_H/\mu_D) = 0.3070122$] quoted in the text are known to better than 0.5 part per million. Show that interactions not taken into account in the simple magnetic-interaction analysis must increase the right-hand side of the equation for the ratio $\delta\tilde{\nu}_H/\delta\tilde{\nu}_D$ by about 1.7×10^{-4} to bring about agreement between the experimental measurements.

4.22. Consider that the nucleus of a hydrogenic atom in an s state can be displaced by an amount very small in comparison with a_0 without altering the electron-charge distribution. If this is so, show that the change in energy of the s state due to a nuclear displacement δr is

$$\delta U = \frac{4}{3} hc R_\infty \frac{Z^4}{n^3} \left(\frac{\delta r}{a_0}\right)^2$$

4.23. Following a suggestion by Welton the Lamb shift may be thought of as described in Sec. 4.4. As is seen in Sec. 6.5 the number of natural modes of oscillation of an electromagnetic field with frequencies in the range $\delta\nu$ near the frequency ν is $8\pi\nu^2 d\nu/c^3$ per unit volume. From Chap. 1 the amplitude of the electric vector E_0 in a radiation field of energy density U is given by $\epsilon_0 E_0^2 = U$. Assuming, in analogy with the harmonic-oscillator discussion of Sec. 5.4, that the lowest energy in each mode of oscillation of the field is $\hbar\nu/2$, show that the electric vector associated with the least possible electromagnetic field, or "free space," is

$$E_0^2 = \frac{4\pi\hbar\nu^3}{\epsilon_0 c^3} d\nu$$

Show that the average square displacement of a free electron under the influence of the totality of these simple harmonic waves is

$$\overline{\delta r^2} = \frac{\alpha\lambda_c^2}{2\pi^3} \int \nu^{-1} d\nu$$

where the integral is taken between certain maximum and minimum limits of the frequency. α is the fine-structure constant, and λ_c is the Compton wavelength. It has been shown by Bethe that the minimum frequency for which the electron may be

considered free for the purpose of this calculation for hydrogen is effectively 17.8 times the frequency corresponding to the ionization energy of the ground state and that the maximum frequency is that corresponding to the Compton wavelength λ_c . Assuming that the change in potential energy between the nucleus and the electron due to this random motion is sufficiently small that the first terms of a Taylor expansion of the potential $V(r)$ in terms of δr may be used to yield the energy change, show from the preceding problem that the energy of binding of an s electron in a hydrogenlike atom is decreased as a result by an amount

$$\delta\nu = \frac{8}{3\pi} R_\infty \alpha^3 \frac{Z^4}{n^3} \ln \frac{\nu_{\max}}{\nu_{\min}}$$

On inserting the limiting values of the frequency given above, show that this is equal to about 1050 Mc or 3.5 wave numbers per meter for $Z = 1$ and $n = 2$.

4.24. Reverting to the ephemeral element positronium, introduced in Prob. 3.18, and knowing the electron and positron have same mass and spin, show qualitatively that the fine structure of its spectrum should resemble that of helium showing ortho and para types. If the magnetic moments associated with the spins of the electron and positron are in the senses to be expected for a spinning charge of the sign concerned, show that no initial Zeeman effect is to be anticipated. Show by the method of Prob. 4.18 or the result of Prob. 4.20 that the magnetic-interaction-energy difference in the ground state ($n = 1, j = 0, 1$) is $\frac{2}{3}h\alpha^2R_\infty$. Assume that finally the electron and positron combine together to annihilate one another with the emission of one or more photons and in conformity with the three conservation laws of energy, momentum, and angular momentum. What would be the minimum number of photons that would have to be emitted if the isolated positronium atom were in an s state and the spins were (a) antiparallel, or (b) parallel? It will be recalled that when a photon is radiated it takes away one unit of angular momentum.

4.25. Write the terms that arise from the following electron configurations on the assumption that the total quantum numbers of the two electrons are not the same: $p \cdot p$, $p \cdot d$, $d \cdot d$, $p \cdot f$, $d \cdot f$, and $f \cdot f$. Show that if an electron, unrestricted by the Pauli principle, is added to a configuration giving rise to certain terms characterized by S , L , and J , the resulting terms will be characterized by S' , L' , and J' , where $S' = S \pm \frac{1}{2}$; $L' = L + l$, $L' = L + l - 1$, etc., to $L' = L - l$; and $J' = L' + S'$, $J' = L' + S' - 1$, etc., to $J' = L' - S'$. The letter l characterizes the angular momentum of the electron added.

4.26. Derive the terms arising from the electron configuration p^3 , and show that the terms arising from the configuration p^n and p^{6-n} are the same for $n = 1$ and $n = 2$.

4.27. Using the table of the lowest electron configurations of the elements, find the low-configuration terms to be expected for La, Hf, Ta, W, and Re.

4.28. Show by analogy with the calculations in the case of helium that to the first approximation the energy necessary to remove one of the two remaining electrons from a Li^+ ion is 71.4 ev.

GENERAL REFERENCES

- Bacher, R. F., and S. Goudsmit: "Atomic Energy States," McGraw-Hill Book Company, Inc., New York, 1932.
 Conference on Spectroscopy, *Rev. Mod. Phys.*, **14**, 57ff. (1942).
 Conference on Spectroscopy, *Rev. Mod. Phys.*, **16**, 152ff. (1944).
 Le Corbeiller, P.: "Matrix Analysis of Electric Networks," Harvard University Press, Cambridge, Mass., 1940.

- Dirac, P. A. M.: "Principles of Quantum Mechanics," chaps. 7-9, Oxford University Press, London, 1944.
- Hartree, D. R.: *Repts. Progr. in Phys.*, **11**, 113 (1947) (structural calculations).
- Houston, W. V.: "Principles of Quantum Mechanics," McGraw-Hill Book Company, Inc., New York, 1951.
- Moore, C. E.: Atomic Energy Levels, *Natl. Bur. Standards Bull.* 467 (1949).
- Mott, N. F., and I. N. Sneddon: "Wave Mechanics and Its Applications," Oxford University Press, London, 1948.
- Pauling, L., and E. B. Wilson: "Introduction to Quantum Mechanics," chaps. 6-8, McGraw-Hill Book Company, Inc., New York, 1935.
- Pringsheim, P.: "Fluorescence and Phosphorescence," Interscience Publishers, Inc., New York, 1949.
- Rubinowicz, A.: *Repts. Progr. in Phys.*, **12**, 233 (1949) (multipole radiation).
- Schiff, L. I.: "Quantum Mechanics," chaps. 6, 7, 9, 11, 12, McGraw-Hill Book Company, Inc., New York, 1949.
- Spectroscopy, *Repts. Progr. in Phys.*, **5**, 210ff, (1938).
- White, H. E.: "Introduction to Atomic Spectra," chaps. 5, 8-10, 18, McGraw-Hill Book Company, Inc., New York, 1934.
- White, H. E.: *Repts. Progr. in Phys.*, **6**, 145 (1939) (Zeeman effect).

CHAPTER 5

MOLECULAR STRUCTURE

5.1. Electron and Nuclear Locations. In the electronic structure surrounding a nucleus the electrons have an average effective spatial distribution, though this cannot be measured with any great precision. The density function uu^* can be calculated precisely in the case of hydrogen. The density functions of more complicated atoms can be calculated by a process of successive approximations leading to consistency between the potential V of Schrödinger's¹ equation and the potential derived from Poisson's equation using euu^* for the charge density. In atoms with a sufficiently great number of electrons to warrant a statistical approach, an approximation can be made fairly easily by the method of Thomas and Fermi (Sec. 6.5). Though precise experimental verification cannot be expected, approximate information on the electron distribution can be obtained by observing the scattering of photons or material particles from such systems if the wavelengths of the scattered particles are less than the dimensions of the atomic system being studied.

The simplest particles available for scattering experiments are (1) photons, (2) electrons, (3) protons, (4) neutrons. In principle, in order to describe such a scattering process one sets up the Hamiltonian or energy equation representing the energy of interaction between the atomic system and particle to be scattered and endeavors to find solutions corresponding to incoming plane waves and outgoing spherical waves from the atomic system. Comparison of theory and experiment then determines the extent to which the assumptions of no atomic distortion, etc., are valid and yields the values of the assumed parameters representing the atomic electron distribution. The specific interaction of neutrons with electrons is negligibly small. The energy of protons having wavelengths of the order of atomic systems is so small that the experiments with them are very difficult to perform. Proton and neutron scattering experiments have yielded more information about the energy of interaction of these particles with atomic nuclei than with electron structures. Electrons and photons, however, exhibit electrostatic or electromagnetic

¹ D. R. Hartree, *Proc. Cambridge Phil. Soc.*, **24**, 189 (1928); V. Fock, *Z. Physik*, **61**, 126 (1930); J. C. Slater, *Phys. Rev.*, **35**, 210 (1930).

interactions with electrons in the range convenient for experimental techniques. It is from the scattering of these particles that most of our knowledge about the electronic probability density distribution in atoms and molecules has been gained.

Electron and photon scattering yield quite similar information. As an illustration, photon scattering will be considered on the basis of the preceding discussions in Secs. 1.8 and 4.6. As indicated in the earlier classical discussion, photons can be scattered by the electron structure without any change in energy or they may be scattered with a loss of energy. In the first case no energy change is induced in the atomic system nor is an electron ejected as a photoelectron, the photon energy and wavelength are unaltered, and the process is known as *coherent scattering* since the incoming and outgoing waves exhibit the phenomenon of interference. In the second case some or all of the photon energy is transferred to the atomic system as energy of excitation or ionization and any excess kinetic energy may be carried away by one or more electrons. In such a process the energy of the photon is decreased, its wavelength is increased by the Compton or Raman process, and it can no longer interfere with the incident radiation. This is known as *incoherent scattering*. The following discussion is concerned with elastic or coherent scattering only.

In the first place the method of Sec. 4.6 may be used to calculate approximately the polarizability of an atom in an electric field. This is the induced dipole moment per unit electric field which is the analogue of the classical elastic constant of binding of an electron by a nucleus. The change in energy due to the perturbation of an atomic electron by an electric field, either static or associated with an electromagnetic wave, would be from Eq. (4.23)

$$W'_{00} = \int u_0^* H' u_0 d\tau$$

Here the subscript 0 is used to indicate the lowest, or normal, energy state. The perturbing energy H' associated with the presence of an electric field E in the x direction is $eE_x x$. This is an odd function of x . Since u_0 and u_0^* are both odd or both even, the integrand is odd in x and the integral vanishes. Thus there is no first-order energy change due to the field E . The wave function of the perturbed atom, which will be written u'' , is by Eq. (4.23)

$$u'' = u_0 + \sum_k \frac{W'_{0k}}{E_0 - E_k} u_k$$

where

$$W'_{0k} = \int u_k^*(eE_x x) u_0 d\tau$$

The charge density, to first-order terms in W' , is then

$$\rho(x,y,z) = e|u''|^2 = e|u_0|^2 + e \sum_k \frac{u_k u_0^* + u_0 u_k^*}{E_0 - E_k} W'_{0k}$$

In the second term the charge density is not in general spherically symmetric; though u_0 may correspond to a symmetrical s electron, u_k may correspond to a p or d electron including angular factors (see Fig. 5.1). The polarization P_x due to the field E_x is then

$$\begin{aligned} P_{x0} &= \int x\rho d\tau = e \int u_0 x u_0^* d\tau + e \sum_k \frac{W'_{0k}}{E_0 - E_k} \int (u_k x u_0^* + u_0 x u_k^*) d\tau \\ &= \frac{2}{E_x} \sum_k \frac{|W'_{0k}|^2}{E_0 - E_k} \end{aligned}$$

since the first term vanishes and the integrals are obviously related to W'_{0k} . The polarization per unit field, or *polarizability*, α of the normal state is

$$\alpha_0 = \frac{2}{E_x^2} \sum_k \frac{|W'_{0k}|^2}{E_0 - E_k} \quad (5.1)$$

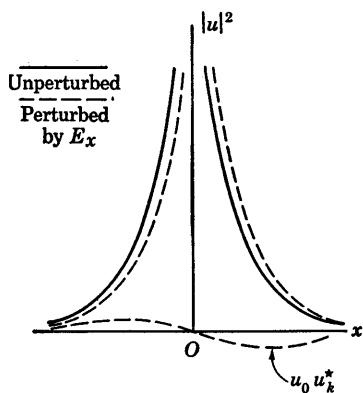


FIG. 5.1. One-dimensional perturbation of an atomic wave function by an electric field.

This is frequently written in terms of a dimensionless quantity known as the *oscillator strength* of the state. This is defined as

$$f_k = \frac{2m}{\hbar^2} (E_0 - E_k) |X'_{k0}|^2 \quad (5.2)$$

where

$$X'_{k0} = \frac{1}{eE_x} \int u_0 (eE_x x) u_k^* d\tau = \int u_0 x u_k^* d\tau$$

In terms of f_k ,

$$\alpha_0 = \frac{e^2}{m} \sum_k \frac{f_k}{\omega_k^2}$$

where $\hbar\omega_k = E_0 - E_k$. $e^2/m\omega^2$, by Sec. 1.8, is the classical value for the polarizability of an oscillator of natural frequency ω . Writing α_{ck} for this quantity, the polarizability becomes

$$\alpha_0 = \sum_k f_k \alpha_{ck} \quad (5.3)$$

It may be shown that $\sum_k f_k = 1$. If hydrogenic wave functions are used to calculate X'_{k0} the polarizability of the hydrogen atom can be obtained.

The above results are of importance in such practical phenomena involving the effect of electric fields on atoms as the Stark effect, Van der

Waals forces, and the scattering of low-frequency radiation (see Sec. 3.6). But these results do not apply to the scattering of radiation having a wavelength comparable to atomic dimensions. In such a case, E_z varies to a significant extent over the dimensions of the atom and cannot be taken as a constant.

The energy associated with a hydrogenic term is seen from Eq. (3.20) to be

$$E_n = -\frac{m_r e^4 Z^2}{2(4\pi\epsilon_0)^2 \hbar^2 n^2} \cong -\frac{\hbar^2}{2m_0} \left(\frac{Z}{na_0}\right)^2$$

where $\alpha_0 = 4\pi\epsilon_0\hbar^2/m_0e^2$, and the electron mass m_0 is written for the reduced mass m_r . From Sec. 3.4, na_0/Z is a measure of the spatial extent of the wave function, and a light wave of this wavelength has an energy $E_l = h\nu = 2\pi\hbar c Z/na_0$. It is thus seen that

$$\frac{E_l}{E_n} = \frac{4\pi m_0 c}{\hbar} \frac{na_0}{Z} = 4\pi \frac{n}{Z\alpha}$$

where α is the fine-structure constant, about $1/137$. E_l is thus about a thousand times E_n for n/Z of the order of unity. If the photon energy is very much greater than the binding energy of the electron, the electron may be thought of as effectively free of atomic binding forces. Thus for X-ray scattering, the classical value of the polarizability, $e^2/m\omega^2$, is generally a satisfactory approximation. Though the electron may be considered to be free in the calculation of scattering, its inclusion in the atom is necessary to ensure conservation of momentum if there is no change in wavelength of the radiation scattered. If the scattering electron were a point, the cross section for scattering at the angle θ by Sec. 1.8 would be $\sigma(\theta) = (r_0 \sin \theta)^2$ and the ratio of the electric vector in the scattered wave to that in the incident wave would be

$$\frac{E_{zs}}{E_{z0}} = (r_0/r) \sin \theta$$

It may be assumed that each portion of the electron, that is, $uu^* dr$, does scatter in this way, and it is then only necessary to allow for the phases of the waves scattered by each infinitesimal volume element as a function of the angle of observation and then to sum over the space occupied by the electron, in order to obtain the scattered amplitude.

Let \mathbf{k} be the propagation vector of the incident wave, \mathbf{r} the vector displacement of the scattering element $uu^* dr$ for each electron from the center of the atom, and \mathbf{k}' the propagation vector of the scattered wave in the direction of observation. The phase of a wave scattered from $uu^* dr$ relative to one scattered from the center of the atom is seen from Fig. 5.2 to be $(2\pi/\lambda)r(\cos \chi - \cos \chi')$, or $\mathbf{r} \cdot (\mathbf{k} - \mathbf{k}')$. Thus the amplitude of the wave relative to unit amplitude scattered from the origin per electron is $uu^* e^{i\mathbf{r} \cdot (\mathbf{k} - \mathbf{k}') dr}$.

The total scattered amplitude then is proportional to the sum over all electrons, or

$$\sum_1^Z \int uu^* e^{ir \cdot s} d\tau \quad \text{where } \mathbf{s} = \mathbf{k} - \mathbf{k}' = \frac{4\pi}{\lambda} \sin \frac{\phi}{2} \mathbf{s}_1$$

Referring again to Sec. 1.8, unit scattering from the origin by an unpolarized beam leads to an intensity cross section (amplitude squared) of $r_0^2(1 + \cos^2 \phi)/2$ so the cross section for coherent scattering by the atom

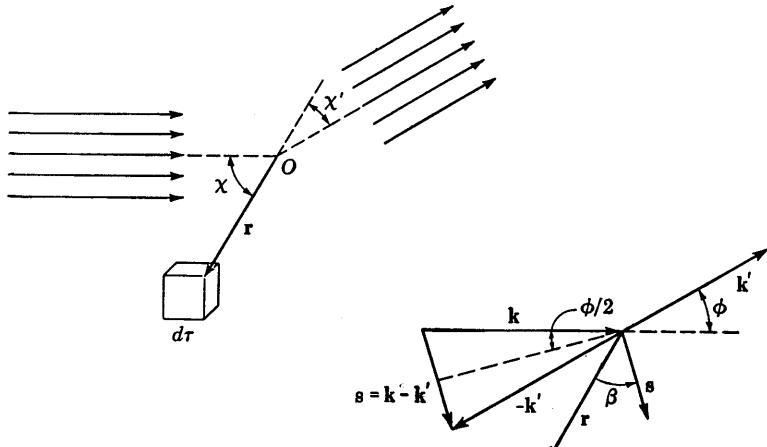


FIG. 5.2. Construction for coherent-scattering analysis.

at an angle ϕ with the incident beam is

$$\sigma(\phi) = r_0^2 \frac{1 + \cos^2 \phi}{2} \left[\sum_1^Z \int uu^* e^{ir \cdot s} d\tau \right]^2 \quad (5.4)$$

In the case of an atom, uu^* may be taken as spherically symmetrical and Eq. (5.4) may be integrated over the angles. If β is the angle between \mathbf{r} and \mathbf{s} (ϕ is a constant), the square bracket, known as the structure factor F , is

$$\begin{aligned} F &= \sum_1^Z \int_0^\infty \int_0^\pi \int_0^\pi uu^* e^{irs \cos \beta} 2\pi r^2 \sin \beta d\beta dr \\ &= \sum_1^Z \int_0^\infty \frac{1}{2irs} [e^{irs} - e^{-irs}] uu^* 4\pi r^2 dr \\ &= \int_0^\infty U(r) \frac{\sin sr}{sr} dr \quad \text{where } U(r) = 4\pi r^2 \sum_1^Z uu^* \end{aligned}$$

Thus if u is known, $\sigma(\phi)$ can be calculated. Alternatively, if F is determined experimentally, $U(r)$ and hence u may be obtained by the Fourier transformation (Sec. 1.4):

$$U(r) = \frac{2}{\pi} \int_0^{\infty} srF(s) \sin sr ds$$

The experimental values by Barrett, Wollan, Herzog,¹ and others on coherent and incoherent scattering agree in close detail with the calculated values of $\sum_1^Z uu^*$ from the Hartree or Thomas-Fermi theory (Sec. 6.5). A typical curve for scattering in argon is shown in Fig. 5.3. For atoms with many electrons, a comparatively simple approximation (Sec. 6.5) in which the form factor is proportional to Z is adequate for most purposes. In Fig. 5.4 this approximation is plotted as a universal function of $\frac{a_0 \sin (\phi/2)}{\lambda Z^{\frac{1}{3}}}$.

a_0 is the Bohr radius, λ is the wavelength, ϕ is the scattering angle, and Z is the atomic number.²

Scattering experiments lead very directly to information on the disposition and spacing of atomic nuclei and electrons in molecular combinations. Making the heuristic assumption that the nuclei and the most closely bound electrons of the constituent atoms retain their identity in molecules and assuming that but relatively few of the outer electrons have their energies significantly altered by molecular combination, one can anticipate the nature of the scattering phenomena that would be observed. If neutrons are scattered from a molecular gas, the scattering or diffraction pattern should be determined by average internuclear distances since the electrons are not influential in such scattering. The diffraction pattern produced by photons or electrons, on the other hand, should be determined primarily by the characteristic separations of major concentrations of electrons in the molecular structure. For heavy elements with many closely bound electrons centered about the nuclei, the two types of scattering should be very similar and the patterns should be alike in their major features. For light elements with relatively few

¹ E. O. Wollan, *Phys. Rev.*, **37**, 862 (1931); G. Herzog, *Z. Physik*, **69**, 207 (1931); C. S. Barrett, *Phys. Rev.*, **32**, 22 (1928).

² P. Debye, *Physik. Z.*, **31**, 419 (1930); L. Bewilogna, *Physik. Z.*, **32**, 740 (1931).

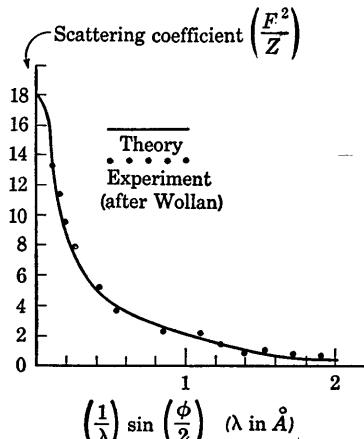


FIG. 5.3. Coherent atomic scattering of X rays by argon.

electrons, and those loosely bound, there would be little variation of electron density over the molecule and little structure to the diffraction pattern. Both types of scattering should be axially symmetrical about the beam because of the random orientation of the scattering molecules in the gas. The above considerations are borne out by experiment. Neutron scattering depends on nuclear properties and internuclear distances; X-ray scattering is determined by the electron distribution.

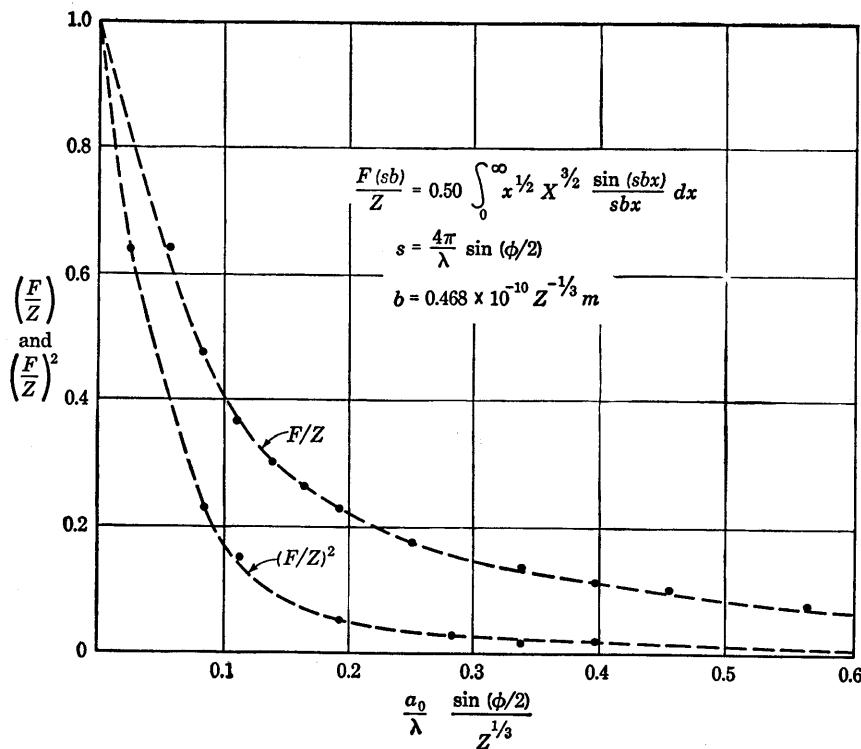


FIG. 5.4. Universal form-factor curves for heavy atoms. (Thomas-Fermi.)

X-ray scattering experiments yield some interpretable information upon this electron distribution, and in the case of heavy atoms internuclear distances may be inferred.

In the case of a polyatomic molecule the amplitude of coherent scattering of X rays is made up from the sum of the scattering by each atomic electron concentration about the nuclei, which, in turn, is approximately proportional to the F value for that atom. Neglecting the variation in F values caused by molecular binding, the scattered X-ray amplitude from a molecule, with due regard to phase, can be obtained by adding up the contribution of each atom. Omitting for the moment the term

$r_0^2(1 + \cos^2 \phi)/2$ and referring to Fig. 5.5, the amplitude of scattering is seen to be proportional to

$$\sum_k F_k e^{i\mathbf{r}_k \cdot \mathbf{s}}$$

where \mathbf{r}_k is the radius vector from some arbitrary origin to the center of atom k , and $\mathbf{s} = (4\pi/\lambda) \sin(\phi/2)\mathbf{s}_1$ of the previous discussion. The intensity of scattering from a molecule is proportional to the amplitude times the complex conjugate, i.e.,

$$\sum_k F_k e^{i\mathbf{r}_k \cdot \mathbf{s}} \sum_j F_j e^{-i\mathbf{r}_j \cdot \mathbf{s}} = \sum_{j,k} F_k F_j e^{-i(\mathbf{r}_k - \mathbf{r}_j) \cdot \mathbf{s}}$$

The quantity $\mathbf{r}_k - \mathbf{r}_j$ may be written as \mathbf{l}_{jk} , the vector separation of atom pairs, and the intensity may be averaged over all possible orientations of the molecules in the gas. Intensities rather than amplitudes are averaged because the relative positions and orientations of molecules are randomly distributed and the scattering from several molecules is not coherent.

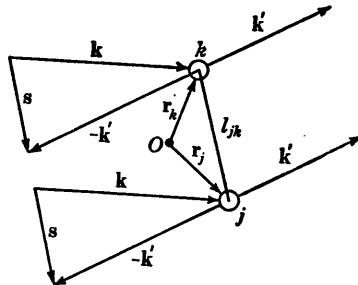


FIG. 5.5. Coherent scattering from two atomic centers.

$$\begin{aligned} \text{Average intensity} &\approx \frac{\int \sum_{j,k} F_j F_k e^{i l_{jk} s \cos \theta} \sin \theta d\theta}{\int \sin \theta d\theta} \\ &\approx \sum_{j,k} F_j F_k \frac{\sin l_{jk} s}{l_{jk} s} \end{aligned}$$

Recalling the factor $r_0^2(1 + \cos^2 \phi)/2$, the cross section for coherent scattering at an angle ϕ due to a molecular gas becomes

$$\sigma(\phi) = r_0^2 \frac{1 + \cos^2 \phi}{2} \sum_{j,k} F_j F_k \frac{\sin l_{jk} s}{l_{jk} s} \quad (5.4')$$

The term $(\sin l_{jk} s)/l_{jk} s$ is a periodic function of $l_{jk} s$, as shown in Fig. 5.6. From the spacing of maxima as a function of ϕ , the interatomic spacings l_{jk} can be determined. In a simple diatomic molecule the sum becomes

$$F_1^2 + F_2^2 + 2F_1 F_2 \frac{\sin l_s}{l_s}$$

For a very light molecule such as H_2 the theory does not apply since there is no electron density near the nucleus and the internuclear distance

for such a light molecule must be obtained from neutron scattering. However, for oxygen, O_2 , with eight electrons associated with each atom, the periodic variation of $\sigma(\phi)$ with angle is observed and the interatomic spacing of 1.22 Å is found by experiment. An illustration of a more

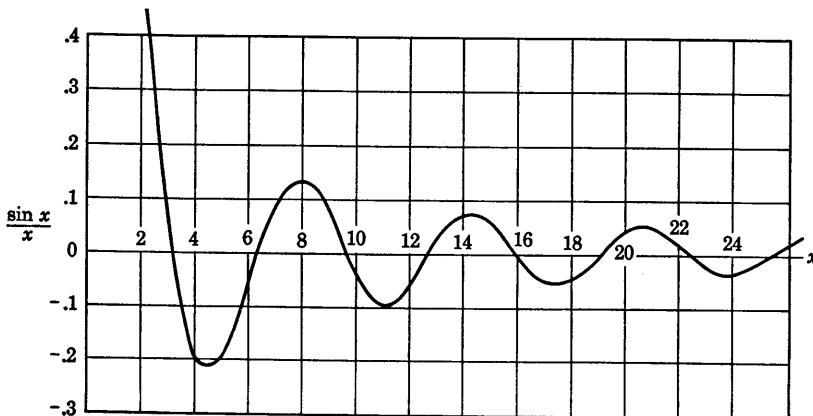


FIG. 5.6. Pairwise interference scattering factor.

complicated molecule is carbon tetrachloride, CCl_4 . The most symmetric arrangement of the atoms is that in which the chlorine atoms are at the corners of a tetrahedron centered on the carbon atom as in Fig. 5.7. The sum term is then

$$F_c^2 + 4F_{Cl}^2 + 12F_{Cl}^2 \frac{\sin l_{1s}}{l_{1s}} + 8F_c F_{Cl} \frac{\sin l_{2s}}{l_{2s}}$$

F_c is much smaller than F_{Cl} , since the carbon atom has but two electrons

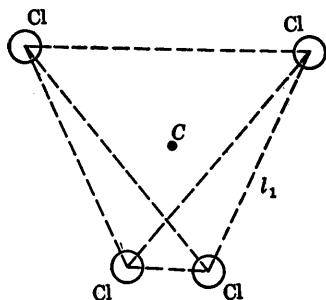


FIG. 5.7. Tetrahedral CCl_4 atomic centers.

presented yields a consistent description of the phenomena observed. Also the results of X-ray scattering experiments are concordant with those from other types of experiments such as chemical evidence, crystal

$$4F_{Cl}^2 \left(1 + 3 \frac{\sin l_{1s}}{l_{1s}} \right)$$

The experimental curves of $\sigma(\phi)$ for this molecule show the simple periodicity associated with one length (l_1), confirming the model, and the separation of the maxima yields a value of 2.99 Å for l_1 . Thus the general but simple theory here

structure studies, and observations of molecular spectra and products of ionization in molecules.¹

5.2. Nature of Molecular Energy Levels. A molecule is a much more complex structure than an atom because it consists of several nuclei and their accompanying electrons and hence is an entity which inherently presents a many-body problem. In consequence, the deductive theory of its structure is necessarily more qualitative and approximate. Following the discussion of Sec. 3.4 the formal expression for the energy, i.e., Schroedinger's equation, can be written down. If the center of gravity of the molecule is assumed to be fixed so that the translational kinetic energy is removed from consideration and the time-dependent factor is taken to be $e^{-iWt/\hbar}$ where W then represents the permissible total energy values, Eq. (3.6) becomes

$$\sum_{j=1}^{j=p} \frac{1}{M_j} \nabla_j^2 u + \frac{1}{m} \sum_{k=1}^{k=q} \nabla_k^2 u + \frac{2}{\hbar^2} (W - V) u = 0 \quad (5.5)$$

Here there are assumed to be p different nuclei of mass represented by M_j , q electrons each of mass m , and V is the potential energy of the system. This represents a very complicated physical situation even if the potential energies considered are limited to those of electrostatic origin and the magnetic interactions of spins and orbital angular momenta are neglected. If this simplest assumption is made, V is given by

$$4\pi\epsilon_0 V = \sum_{k,k'} \frac{e^2}{r_{kk'}} + \sum_{j,j'} \frac{Z_j Z_{j'} e^2}{r_{jj'}} - \sum_{j,k} \frac{Z_j e^2}{r_{jk}}$$

Ordinary perturbation methods cannot be applied directly because no solutions of an equation simpler than but comparable to Eq. (5.5) are known.

The best approximate method of dealing with Eq. (5.5), which has been justified theoretically by Born and Oppenheimer² and which yields results in conformity with experimental observation, is essentially to separate it into those terms which primarily involve the outer electron structure of the constituent atoms and those which primarily involve the motion of the nuclei and their closely bound electron shells. This receives qualitative physical justification by considering the motion of the outer loosely bound electrons and the heavier nuclei together with their associated inner electron shells under the influence of the forces of interaction between them. The relatively light outer electrons may be thought of

¹ E. O. Wollan, *Rev. Mod. Phys.*, **4**, 206 (1932); J. H. McMillen, *Rev. Mod. Phys.*, **11**, 84 (1939); N. S. Gingrich, *Rev. Mod. Phys.*, **15**, 90 (1943).

² M. Born and J. R. Oppenheimer, *Ann. Physik*, **84**, 457 (1927).

as moving much more rapidly under these forces than do the heavier constituents; hence to a first approximation the latter may be thought of as fixed in space in considering the electron motion. From this point of view the electrons complete many cycles of their motion during an interval in which the nuclei move but little. The slower vibrational nuclear motion may then be thought of as taking place within the average configuration of these outer electrons. If x and n are the coordinates and the quantum numbers, respectively, for the electrons, and ξ and ν the like quantities for the nuclei, the assumption is that $u_{n\nu}(x, \xi)$, which is a solution of Eq. (5.5), can be written as $u_n(x, \xi)u_{n\nu}(\xi)$, where $u_n(x, \xi)$ is the electronic function with quantum numbers n depending on the nuclear and electronic coordinates and $u_{n\nu}(\xi)$ is the nuclear function depending on the electron configuration only through the quantum-number parameters specifying it and not upon the electron coordinates. Neglecting terms in $1/M$ in comparison with those in $1/m$ for the zeroth approximation, Eq. (5.5) becomes

$$\sum_{k=1}^{k=q} \nabla_k^2 u_n(x, \xi) + \frac{2m}{\hbar^2} [U_n(\xi) - V(x, \xi)] u_n(x, \xi) = 0 \quad (5.5')$$

where $U_n(\xi)$ is the portion of W associated with electron motion and V is noted as depending upon the coordinates of both electrons and nuclei. Having solved this equation for $u_n(x, \xi)$ and determined the permissible values of $U_n(\xi)$, these latter may be used to determine the characteristic allowed energy values for the first-order approximation in which terms of $1/M$ magnitude are included to determine the slower nuclear motion. Subtracting the zeroth-order equation from the complete equation, which is accomplished by multiplying the terms in Eq. (5.5') by $(1/m)u_{n\nu}(\xi)$ and subtracting them from those of Eq. (5.5), yields

$$\sum_{j=1}^{j=p} \frac{1}{M_j} \nabla_j^2 [u_n(x, \xi)u_{n\nu}(\xi)] + \frac{2}{\hbar^2} [W_{n\nu} - U_n(\xi)] [u_n(x, \xi)u_{n\nu}(\xi)] = 0$$

where subscripts are written on W to indicate that this quantity will depend on the parameters n and ν . Usually the electron configuration described by $u_n(x, \xi)$ is a very slowly varying function of the relative motion of the nuclei, and, to the approximation being considered, this function may be taken as constant and removed from behind the operator ∇_j^2 yielding the following equation for the nuclear motion:

$$\sum_{j=1}^{j=p} \frac{1}{M_j} \nabla_j^2 u_{n\nu}(\xi) + \frac{2}{\hbar^2} [W_{n\nu} - U_n(\xi)] u_{n\nu}(\xi) = 0 \quad (5.5'')$$

This equation may then be solved for the eigenfunctions $u_{n\nu}(\xi)$ and the energy eigenvalues $W_{n\nu}$.

Actually the above procedure is too difficult to carry out (except in H_2), but in principle it indicates a feasible method of approach. The energies associated with the electron configurations are the principal energies in order of magnitude, and they resemble the energies associated with atomic electron configurations since V is of the same nature and order of magnitude as the potential energy of electrons in atoms. The energies associated with nuclear motion can be seen in the following way to be considerably smaller. The frequencies associated with the relative motion of the nuclei are those implied by the forces operative on either electrons or nuclei and the masses of the nuclei. The force equation in order of magnitude would be

$$\frac{m\ddot{x}}{x} = -k = \frac{M\ddot{\xi}}{\xi}$$

where k is the elastic constant. For harmonic motion of the form $e^{i\omega_j t}$ for nuclei and $e^{i\omega_e t}$ for electrons, the ratio of ω_e to ω_n is seen to be of the order of $(m/M)^{\frac{1}{2}}$. Thus the energies associated with the relative motion of nuclei are less in order of magnitude than those associated with electron motion by this factor. Actually, since the molecule is an extended elastic body it can rotate about an axis through its center of gravity without any relative motion of the nuclei, and energy will be associated with this motion as well. In order of magnitude the energies associated with electronic orbital angular momenta in atoms are the same as other electronic energies. If the rotation of atomic nuclei have angular momenta of the order of \hbar , the energies associated with such motion are less than the electronic energies by the factor m/M . For since

$$ma_e^2\omega_e \approx \hbar \approx Ma_n^2\omega_n$$

where the separations of the electrons a_e and of the nuclei a_n are of the same order, the ratio $\hbar\omega_e/\hbar\omega_n$ is of the order of M/m . Thus the orders of magnitude of the energies associated with electronic motion, nuclear relative or vibrational motion, and nuclear rotational motion are in the following relation to one another:

$$\begin{array}{ccc} E_{\text{electronic}} & E_{\text{nuclear vibration}} & E_{\text{nuclear rotation}} \\ 1 & : (m/M)^{\frac{1}{2}} & : m/M \end{array}$$

These three types of motion may then be considered as successive relatively independent approximations to actual molecular energy levels.

The above discussion serves to clarify the general phenomena observed in molecular spectra. Almost all types of molecules emit light in the ultraviolet and visible regions, which corresponds to energy changes of

from 2 to 10 ev. These spectra are of highly varied types for different molecules, but in general they are closely and regularly spaced groups of lines together with associated regions of continua. The emission of radiation in this spectrum region indicates that electronic transitions are taking place, as this is the region in which atomic spectra are observed. The first-order fine structure of these groups of lines is such that $\Delta\lambda/\lambda$ or $\Delta\nu/\nu$ is of the order of 10^{-2} which agrees qualitatively with $(m/M)^{\frac{1}{2}}$, and hence this structure is reasonably associated with nuclear vibratory motion. The radiation associated with pure vibratory motion for which there is no change in the electronic configuration would be in the near infrared, i.e., the wavelength region of from 1 to 10μ . Molecular spectra of this type are observed. Finally, a second-order fine structure is observed in molecular bands in the visible region for which $\Delta\lambda/\lambda$ or $\Delta\nu/\nu$ is again of the order of 10^{-2} of the first-order fine structure. This same order of energy separation is observed as a first-order fine structure in the near-infrared vibrational spectra, and finally in the very far infrared these low frequencies are observed as fundamental emission or absorption bands. These frequencies are clearly to be associated with rotational motion, and detailed analysis confirms this interpretation.

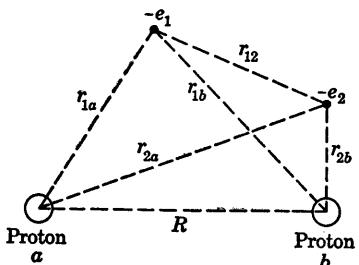


FIG. 5.8. Construction for hydrogen-molecule energy calculation.

The hydrogen molecule is used for illustration. In accordance with Fig. 5.8 the nuclear coordinates ξ reduce to R which is the only nuclear-separation coordinate. The vectors r_1 and r_2 represent the electron coordinates x of Eq. (5.5'). Equation (5.5') then becomes

$$\left\{ \nabla_1^2 + \nabla_2^2 + \frac{2m}{\hbar^2} \left[U_n(R) - \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{R} + \frac{1}{r_{12}} - \frac{1}{r_{1a}} - \frac{1}{r_{1b}} - \frac{1}{r_{2a}} - \frac{1}{r_{2b}} \right) \right] \right\} u_n(r_1, r_2) = 0 \quad (5.6)$$

As a first approximation it is reasonable to try combinations of hydrogenic wave functions as solutions. These may be either symmetric or antisymmetric in the coordinates, the spin choice rendering the total function antisymmetric. It may be shown that such wave functions lead at least to an upper limit for the energy $U(R)$. Making such a

choice

$$u_n(r_1, r_2) = u_a(r_{1a})u_b(r_{2b}) \pm u_b(r_{1b})u_a(r_{2a})$$

where $u_a(r_{1a})$ means the wave function for electron 1 as a component of the hydrogen atom containing proton a , and analogously for the other subscripts. The total electronic characteristic energy values are obtained by multiplying Eq. (5.6) in the abbreviated form $[H - U(R)]u_n = 0$ by u_n^* and integrating over the coordinates of both electrons.

$$\iint u_n^* H u_n d\tau_1 d\tau_2 - \iint u_n^* U(R) u_n d\tau_1 d\tau_2 = 0$$

or, as $U(R)$ is not a function of r_1 or r_2 but only of R ,

$$U(R) \cong \frac{\iint u_n^* H u_n d\tau_1 d\tau_2}{\iint u_n^* u_n d\tau_1 d\tau_2}$$

to this approximation. Using the expression for $u_n(r_1, r_2)$ above,

$$U(R) = \frac{H_{11} \pm H_{12}}{1 \pm \Delta^2} \quad (5.7)$$

where

$$\begin{aligned} H_{11} &= \iint u_a(r_{1a})u_b(r_{2b}) H u_a^*(r_{1a})u_b^*(r_{2b}) d\tau_1 d\tau_2 \\ H_{12} &= \iint u_a(r_{1a})u_b(r_{2b}) H u_a^*(r_{2a})u_b^*(r_{1b}) d\tau_1 d\tau_2 \\ \Delta^2 &= \iint u_a(r_{1a})u_b(r_{2b}) u_a^*(r_{2a})u_b^*(r_{1b}) d\tau_1 d\tau_2 \end{aligned}$$

and the plus sign corresponds to a molecular singlet state (spins anti-parallel) and the minus sign corresponds to a molecular triplet state (spins parallel). Making use of the fact that the u 's are solutions of an equation of the form

$$\left(-\frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{4\pi\epsilon_0 r} \right) u = W_H u$$

where W_H represents the atomic hydrogen energy eigenvalues, the above integrals can be evaluated for the ground states of hydrogen. They are found to be¹

$$\begin{aligned} U(R)_{\text{singlet}} &= 2W_{H_0} + \frac{e^2}{4\pi\epsilon_0 R} + \frac{2J + J' + 2\Delta K + K'}{1 + \Delta^2} \\ U(R)_{\text{triplet}} &= 2W_{H_0} + \frac{e^2}{4\pi\epsilon_0 R} + \frac{2J + J' - 2\Delta K - K'}{1 - \Delta^2} \end{aligned} \quad (5.8)$$

¹ See for instance L. Pauling and E. B. Wilson, "Introduction to Quantum Mechanics," McGraw-Hill Book Company, Inc., New York, 1935.

where

$$\begin{aligned} J &= -\frac{e^2}{4\pi\epsilon_0} \int u_{a0}(r_{1a}) \frac{1}{r_{1b}} u_{a0}^*(r_{1a}) d\tau_1 \\ J' &= \frac{e^2}{4\pi\epsilon_0} \iint \frac{[u_{a0}(r_{1a})u_{b0}^*(r_{2b})]^2}{r_{12}} d\tau_1 d\tau_2 \\ K &= -\frac{e^2}{4\pi\epsilon_0} \int u_{b0}(r_{1b}) \frac{1}{r_{1b}} u_{a0}^*(r_{1a}) d\tau_1 \\ K' &= \frac{e^2}{4\pi\epsilon_0} \iint \frac{u_{a0}(r_{1a})u_{b0}(r_{2b})u_{a0}^*(r_{2a})u_{b0}^*(r_{1b})}{r_{12}} d\tau_1 d\tau_2 \\ \Delta &= \int u_{a0}(r_{1a})u_{b0}^*(r_{1b}) d\tau_1 \end{aligned}$$

The subscript zero indicates the ground state. The values of the integrals can be found from tables or by consulting the reference just given.

Figure 5.9 gives the curves of U from Eq. (5.8) as a function of R/a_0 where a_0 is the Bohr radius, U being measured in rydbergs (R_∞). It is interesting to note that the singlet state, for which the spins are antiparallel, is the lower. This is in distinction to the case of the two electrons in an excited helium atom, where the triplet state, for which the spins are parallel, is the lower. This may be understood qualitatively in the following way. If two electrons are to occupy the same spatial position, their spins must be oppositely oriented by the Pauli principle. Thus spatial contiguity of two electrons may be thought of as favoring a singlet state, for which the spins are antiparallel. In the helium atom the electrostatic repulsion between the electrons tends to separate them, favoring the triplet state. In the hydrogen molecule the attractive electrostatic forces between the two nuclei and the electrons tend to concentrate the electrons between the nuclei, forcing them together and thus favoring the singlet state.

At great separations the energy of the two atoms is simply the sum of the energy of each. As the two atoms are brought closer, moving to the left in Fig. 5.9, the rising triplet-state energy indicates repulsion for parallel spins and the falling singlet state indicates attraction for antiparallel spins till the value of about $1.5a_0$ is reached. Thereafter both states are repulsive. The minimum of the singlet-state curve represents the stable nuclear separation of the molecule of 0.80 \AA as compared with the experimentally observed value of 0.74 \AA . The energy of molecular dissociation is approximately the depth of the minimum below $-2R_\infty$. This is seen from the figure to be about $0.23R_\infty$, which is about 3.14 ev compared with the experimental value of 4.77 ev. The curvature of $U(R)$ near the minimum is within about 10 per cent of the experimental value. Considering the complexity of the problem and the crudity

of the approximations, the agreement is very satisfactory. More refined approximations by James and Coolidge¹ lead to a value of these parameters agreeing with experiment to within the errors of the latter.

Molecular Binding. Of more general importance is the qualitative understanding of molecular binding forces that the foregoing analysis yields. The electron exchange energy, the *K* type of integrals of earlier paragraphs, is capable of yielding stable atomic aggregates essentially through the sharing of pairs of electrons with oppositely oriented spins.

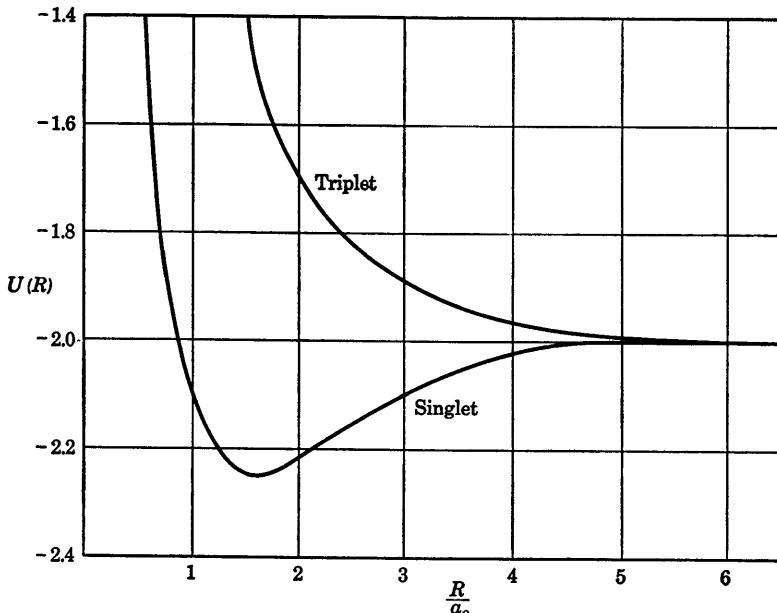


FIG. 5.9. Ground states of the hydrogen molecule (calculated).

This is known as the *covalent bond* or as *homopolar binding*. The bond energy is of the same order as that with which an outer electron is held in an atomic structure. This is the mechanism for holding together stable molecules of the type such as H₂, O₂, or N₂ for which there are no obvious classical attractive forces.

Other types of molecules, such as HCl, may also be considered from the same point of view, but the asymmetry of the constituent atoms is such that the paired electrons (1s from hydrogen and 3p from chlorine) are practically integral parts of the chlorine structure. This results in a molecule which is essentially a single proton associated with an electron structure characteristic of argon surrounding a nucleus lacking one unit of positive charge. If the two components are separated beyond the equi-

¹ H. M. James and A. S. Coolidge, *J. Chem. Phys.*, **6**, 730 (1938).

librium position, they tend to come apart as the proton H^+ and the negative chlorine atom Cl^- . Thus this type of molecular binding is sometimes called *ionic* or *heteropolar binding*, and a simple consideration of the Coulomb forces between point charges gives a good approximation for the forces between constituents at distances somewhat greater than the equilibrium separation.

Except for the form of the potential-energy term the electronic Schroedinger equation for a molecule is of the same form as that for an atom. In consequence, it would be anticipated that the electronic spectroscopic terms would bear a close resemblance in the two cases. This is in fact true, and the only qualitative differences between atomic and molecular energy levels are due to the differences in the symmetry conditions for atomic and molecular systems and to the presence of nuclear separations as parameters in molecular systems. Atomic energy levels are always determined by spherically symmetric potential fields, whereas molecules have their symmetry determined by the relative positions of the atomic nuclei. In the case of a diatomic molecule or linear polyatomic molecule the symmetry is at the least of a cylindrical type. Reflexion of the wave function in a plane containing the axis, i.e., passing through the two nuclei, must result in a wave function of the same magnitude. It may of course change sign since only $|u|^2$ is physically significant; if it changes sign the designation of the term may bear a superscript $(-)$, if it does not the superscript may be $(+)$. If the diatomic molecule is composed of the same type of atom or the linear polyatomic molecule has equal numbers of the same atomic types symmetrically disposed about a center, this point is a center of symmetry. As in the case of an atom the wave functions are then equal in magnitude and of the same or different sign for reflexion in the origin, i.e., a change of sign of the three spatial coordinates with the origin at this center. These two possibilities are known as even or odd, respectively, and designated customarily by the subscripts g (gerade) and u (ungerade). In general polyatomic molecules the symmetry is lower, and for an account of it the special references should be consulted.

There is a close relationship between the electronic angular momenta of the separate constituent atoms and of the molecule they form, but its general detailed analysis becomes quite complex. Consideration of the simple diatomic molecule is illustrative of the more general case, for to a crude approximation polyatomic molecular bonds may be considered as simply so many separate diatomic bonds. A number of the qualitative features of polyatomic molecular structure emerge from a consideration of diatomic molecules. The axial symmetry and intense electrostatic fields of the nuclei determine the nature of the simpler phenomena. If neither one of the electrons forming the bond has orbital angular momen-

tum, as for instance the s electrons in the ground state of H_2 , the resulting electronic structure of the molecule has none, but if one or both of the electrons have orbital angular momentum, i.e., are p or d electrons, the molecule will in general have angular momentum also. The axial electric field tends to polarize each atom. This leads to a separation of the centers of positive and negative charge as seen earlier in this chapter. If the electron has orbital angular momentum l in an arbitrary direction, a torque will develop because of the axially symmetrical electric field acting upon the electron tending to make it precess about the axis of symmetry as in Fig. 5.10. This is analogous to the precession of l in a magnetic field considered in earlier chapters, except that to a first approximation the sense of l is not relevant to the energy in this case. In the magnetic situation the energy is equal to the scalar product of the moment and the magnetic induction; in the electric case it is the polarizability times the square of the electric field and hence always of one sign, though when magnetic effects due to nuclear motion are considered the two orientations of l do not have identical energy values. As in the case of axial symmetry brought about by the existence of a magnetic induction, the component of l along the axis is quantized and can take only integral values of \hbar , but it is only the magnitude of this component that is significant and this is customarily written Λ or $\lambda\hbar$. Molecules for which Λ has the values 0, 1, 2, etc., times \hbar are said to be in Σ , Π , Δ , etc., states in analogy with S , P , D , etc., for atoms. If there is a net electron spin of the configuration it also contributes to the total electronic angular momentum. For a further discussion of the angular-momentum properties of electronic states and transitions between them the references should be consulted.

The calculation of the electronic energy as a function of internuclear distance in any but the simplest cases presents problems that have so far proved insurmountable. The general features of (a) a rapid rise for close approach, (b) a minimum in the case of a stable configuration, and (c) the asymptotic approach to the sum of the atomic energies for increasing R are required of any empirical expression. Several forms have been proposed and used, one of which, composed of two exponential functions, is shown in Fig. 5.11. Its shape in the neighborhood of the minimum represents the elastic characteristics of the electron structure holding the nuclei in position and is a factor in determining the nuclear motions discussed in Sec. 5.4. More than one of these potential curves is of course associated with each molecule, for one asymptotic pair of atoms

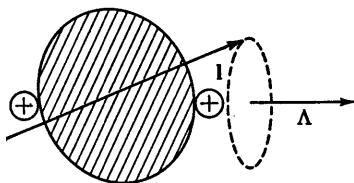


FIG. 5.10. Schematic depiction of electronic angular momentum of a hydrogen-molecule ion.

may give rise to several curves and various possible asymptotic possibilities such as s , p , d , etc., states exist. Thus there is a plurality of molecular electronic energy states, just as there are many electronic energy levels for an atom. There is also the possibility that some potential-energy curves arise from the association of ions instead of neutral atoms to form the normal molecule. The implications of the shapes of these curves and their possible intersections will be mentioned briefly later.

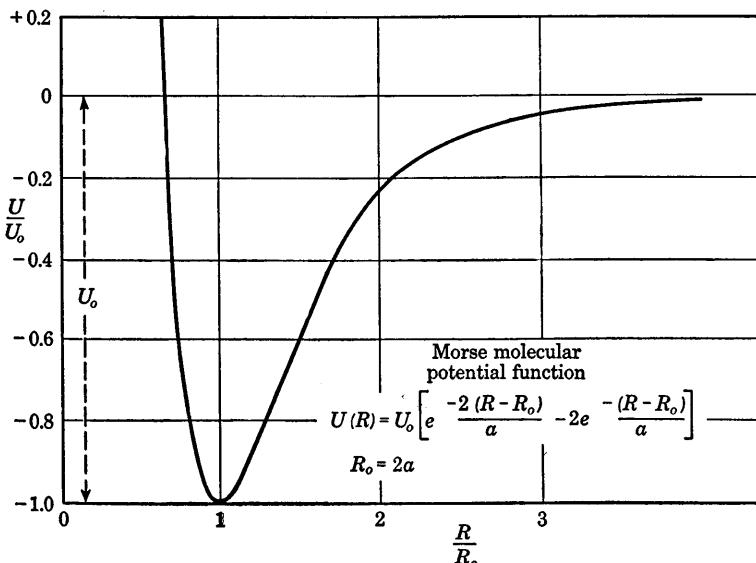
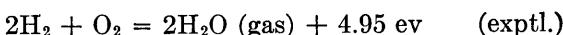


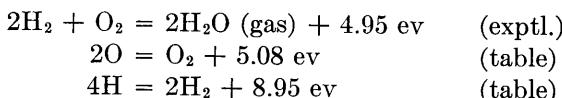
FIG. 5.11. Empirical analytical form of a diatomic-potential-energy curve.

The energy difference between the minimum and the asymptote representing the widely separated constituents is approximately the binding energy U which is of primary importance in chemistry. To a first approximation polyatomic bonds may be considered as the sum of the component diatomic bonds, and a knowledge of the latter is very enlightening in an understanding of chemical reactions. The formation of water from hydrogen and oxygen takes place with explosive violence but requires to be triggered as by a spark. It does not take place simply upon the collision of hydrogen and oxygen molecules. The over-all chemical reaction is



The energy liberated in a chemical reaction is generally measured in kilocalories per mole, and the reaction is assumed to take place at constant pressure. In the above equation and those that follow, the energy

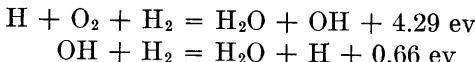
U is expressed in electron volts per molecule and constant-volume processes are assumed so that no correction need be made for the work of expansion of the reacting gases against the surrounding atmosphere. OH molecules, known chemically as *free radicals*, have a temporary existence and are found spectroscopically in a hydrogen-oxygen explosion. Using the values of $U(H_2)$ and $U(O_2)$ determined spectroscopically (Table 5.1), the energy of dissociation of H_2O may be found:



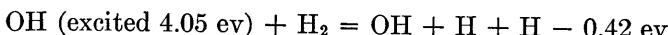
Adding these equations and dividing by 2,



Thus 9.49 ev is the complete dissociation energy of a water molecule. It is somewhat greater than $2U(OH)$ because of the interaction of the two hydrogen bonds. Guided by the knowledge that OH is present, one is led to the conclusion that the following processes, which are energetically possible on the basis of internal energy alone, start off the water-forming reaction:



Thus if hydrogen atoms or OH free radicals are formed by a spark or photodissociation, these chain reactions may proceed with the formation of water. Actually the excess energy of the first reaction is sufficient to elevate the OH radical to an excited electronic state of 4.05 ev, leaving only about 0.24 ev of kinetic energy instead of the 4.29 ev of the equation. In the excited state certain additional reactions which are endothermic but require little more than 4.05 ev can take place at high temperatures. These initiate branching chain reactions since they result in more than one product capable of reacting with H_2 and O_2 molecules. An example is the simple dissociation of a hydrogen molecule on encountering an excited radical:



Thus if about 0.5 ev of kinetic energy is available, several additional free radicals are produced per unit process and the reaction can proceed at an ever increasing rate resulting in an explosion. In a similar way the reaction $H_2 + Cl_2 = 2HCl$ can be induced to take place in the stages $Cl + H_2 = HCl + H$, $H + Cl_2 = HCl + Cl$, etc., upon, say, the photo-dissociation of Cl which from the table evidently requires 2.476 ev. This

corresponds to a wave number of $2.476e/hc = 20,000 \text{ cm}^{-1}$, or a wavelength less than 5000 Å.

TABLE 5.1. DIATOMIC MOLECULAR CONSTANTS (after Herzberg and Hagstrum)†

Molecule	$U_0 - \frac{1}{2}\hbar\nu_v, \text{ ev}^\ddagger$	$R_0, \text{ Å}$	$a, \text{ Å}$	$\bar{\nu}_r, \text{ cm}^{-1}$	$\bar{\nu}_v, \text{ cm}^{-1}$
H ₂ ⁺	2.648	1.06	0.731	59.6	2297
H ₂	4.476	0.7416	0.495	121.62	4395.2
HD	4.511	0.7413	0.427	91.31	3817.1
D ₂	4.553	0.7416	0.349	60.86	3118.5
CH	3.47	1.120	0.492	28.91	2861.6
CD	3.52	1.119	0.363	15.62	2101.0
NH	(3.8)	1.038	0.445	33.30	(3300)
OH	4.35	0.971	0.419	37.74	3735.21
HCl	4.430	1.275	0.524	21.182	2989.74
O ₂	5.080	1.2074	0.349	2.8912	1580.36
NO	5.30	1.151	0.322	3.409	1904.03
N ₂	7.373	1.116	0.347	3.849	2207.19
CO	9.60	1.1281	0.389	3.863	2170.21
C ₂	(3.6)	1.312	0.345	3.265	1641.35
Cl ₂	2.475	1.988	0.486	0.4876	564.9
Br ₂	1.971	2.283	0.499	0.1618	323.2
I ₂	1.5417	2.666	0.529	0.0747	214.57
Li ₂	1.03	2.672	1.13	1.3454	351.43
K ₂	0.514	3.923	1.28	0.1124	92.64
Cs ₂	0.45				41.990

$$U(R) = U_0 [e^{-2(R-R_0)/a} - 2e^{-(R-R_0)/a}] = U_0 [e^{-(R-R_0)/a} - 1]^2 - U_0$$

$$\bar{\nu}_v = \frac{1}{2\pi c} \sqrt{\frac{-2U_0}{a^2 M_r}}, \quad \bar{\nu}_r = \frac{\hbar}{2\pi c M_r R_0^2}$$

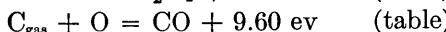
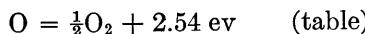
† G. Herzberg, "Spectra of Diatomic Molecules," D. Van Nostrand Company, Inc., New York, 1950; H. D. Hagstrum, *Rev. Mod. Phys.*, **23**, 185 (1951).

‡ 1 ev = 23.06 kg-cal mole⁻¹.

Organic Compounds. Calculations of the above type relating to organic compounds involve the additional consideration that under the ordinary conditions of pressure and temperature carbon is not a gas but a solid. Thus the heat of sublimation of carbon is an important quantity. The burning of solid carbon yields 26.8 cal mole⁻¹, or



also



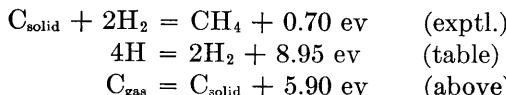
Thus the latent heat of carbon is given by

$$L = C_{\text{gas}} - C_{\text{solid}}$$

or

$$L = O + C_{\text{gas}} - CO - 3.70 = 9.60 - 3.70 = 5.90 \text{ ev}$$

Using this value and the heat of formation of methane by burning carbon in hydrogen, the dissociation energy of methane can be found:



Therefore

$$4H + C_{\text{gas}} = CH_4 + 15.55 \text{ ev}$$

This is greater by about 12 per cent than four CH bonds, the discrepancy being due to the neglect of the interactions of the hydrogen atoms in the molecule. The two illustrations given above are representative of the application of band-spectrum information to chemistry. On the assumption of independence of the bonds in a molecule, the energies of quite complex molecules can be calculated approximately by considering these energies to be simply the sums of the energies of the individual bonds.

Polyatomic molecules cannot be described by the single separation parameter R but require a number of such parameters sufficient to locate all the constituent atomic nuclei. It is found experimentally that the angles between the vector separations of the atoms neighboring a particular atom in a molecule are characteristic quantities; some atoms like CO_2 are linear (the angle between the vectors joining the C to the two O's is 180°), whereas others such as H_2O are not. The angle between the vectors joining the O to the two H's is about 105° . Thus the electron structure in which the atomic nuclei are imbedded to form the molecule may be thought of as an elastic structure with resistance to bending as well as to extension or compression. Though quantitative calculations are very difficult and their accuracy is not rewarding, the general nature of the phenomena can be described and partially understood in terms of the hydrogenic wave functions of Sec. 3.4. There it was seen that s wave functions are spherically symmetrical; u is a function of r only, or $u(s) = f_s(r)$. The p wave functions, from Sec. 3.4, have an angular dependence such that

$$u(p_x) = \frac{x}{r} f_p(r), \quad u(p_y) = \frac{y}{r} f_p(r), \quad u(p_z) = \frac{z}{r} f_p(r)$$

In accordance with the general discussion earlier in this chapter, the binding energy is determined by integrals of the form $\int \frac{u_1 u_2}{r} d\tau$ which are clearly dependent upon the amount and nature of the overlapping of the wave functions of the two atoms in the volume over which the

integration is performed. Thus, proceeding by analogy with the hydrogen molecule, it might well be assumed that the general features of molecular geometry could be derived by requiring as a criterion for maximum binding energy that the wave functions of electrons unpaired in spin with others in the same atom overlap as much as possible with such unpaired electrons in other atoms making up the molecule. The spin of these bonding pairs would be oriented oppositely as this is the condition under which electrons tend to be closely associated with one another.

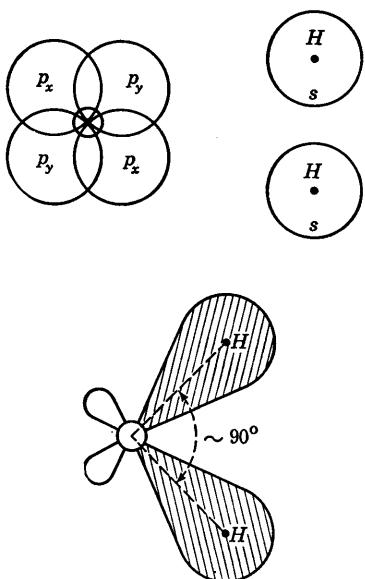


FIG. 5.12. Schematic of orbital overlapping for oxygen and hydrogen atoms in the formation of an H_2O molecule.

Water is a simple illustration. The oxygen atom has the electron configuration $1s^2 2s^2 2p^4$. Oxygen, being divalent, has two of its p electrons internally paired, say in the $u(p_z)$ orbital wave function, and the other two evidently unpaired in $u(p_x)$ and $u(p_y)$, respectively. Normal hydrogen, having a $1s$ electron configuration, has wave functions that are spherically symmetrical. On approaching the oxygen atom along the axes of the unpaired electrons, these oxygen electrons tend to concentrate in that lobe nearer the hydrogen at the expense of the oppositely directed lobe, and the two shared electron bonds thus formed make an angle of about 90° with one another, as indicated in Fig. 5.12. Because of repulsion of the like-oriented spins of the electrons in the two H atoms, these are mutually repelled, giving the observed angle, 105° .

Another illustration is carbon, which normally has the configuration $1s^2 2s^2 2p^2$. As such it is divalent, forming the compound CO , for instance, by means of a double bond with the two electrons of O as in Fig. 5.13. In many cases, however, it is tetravalent, i.e., there are four unpaired electrons. This configuration results from the acquisition of additional angular momentum by one of the $2s$ electrons so that it becomes a $2p$ electron, and the configuration is $1s^2 2s^2 2p^3$. Here there are four electrons to be paired if one of the p 's is in each of the three spatial u functions. A linear combination of the s and three p functions can be made to yield a function of high spatial symmetry having four axes, one for each electron. In fact, this can be done in more than one way: (a) the lobes may make equal angles with one another, extending toward the apexes of a regular

tetrahedron centered on the nucleus, or (b) three may be coplanar at angles of 120° with one another, and the fourth normal to the plane. These relationships may be set up by writing the linear combinations as

$$\begin{aligned} u_1 &= a_1u(s) + b_1[l_1u(p_x) + m_1u(p_y) + n_1u(p_z)] \\ u_2 &= a_2u(s) + b_2[l_2u(p_x) + m_2u(p_y) + n_2u(p_z)] \\ u_3 &= a_3u(s) + b_3[l_3u(p_x) + m_3u(p_y) + n_3u(p_z)] \\ u_4 &= a_4u(s) + b_4[l_4u(p_x) + m_4u(p_y) + n_4u(p_z)] \end{aligned}$$

where the a 's, b 's, l 's, m 's, and n 's are to be determined. Since $u(p_x)$ has its maximum in the direction x , etc., u_1 is a maximum in the direction

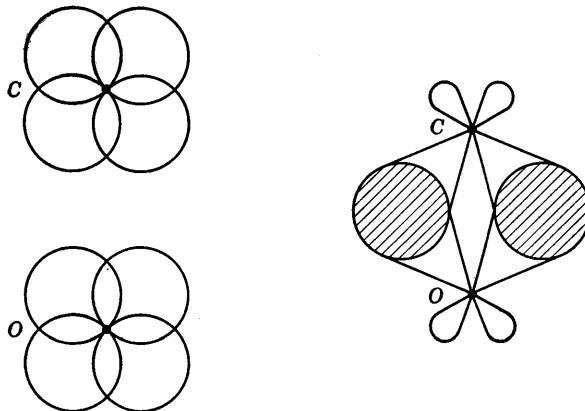


FIG. 5.13. Schematic of orbital overlapping for divalent carbon and oxygen in the formation of CO.

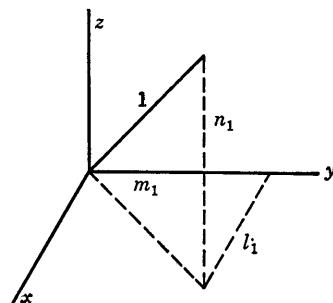


FIG. 5.14. Construction for orbital direction cosines.

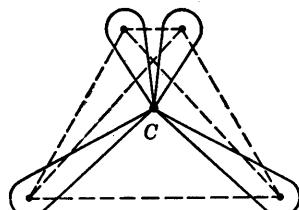


FIG. 5.15. Schematic of tetragonal carbon orbitals.

having the direction cosines l_1 , m_1 , and n_1 as in Fig. 5.14. Normalization, if desired, is seen to require

$$a_i a_j + b_i b_j (l_i l_j + m_i m_j + n_i n_j) = \delta_{ij}$$

or

$$a_i a_j + b_i b_j \cos \theta_{ij} = \delta_{ij}$$

where θ_{ij} is the angle between the maxima of u_i and u_j . For the tetrahedral case (*a*) above, the bonds may be thought of as in the directions with cosines proportional to $(1,1,1)$, $(-1,-1,1)$, $(1,-1,-1)$, and $(-1,1,-1)$; the a 's and b 's are determined by the normalization condition; and the u_j 's may be written down at once. A schematic diagram of the tetrahedral arrangement is given in Fig. 5.15. For the trigonal case

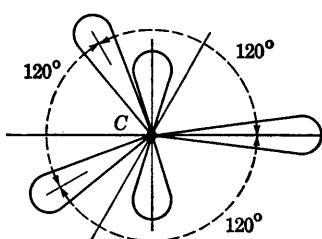


FIG. 5.16. Schematic of trigonal carbon orbitals.

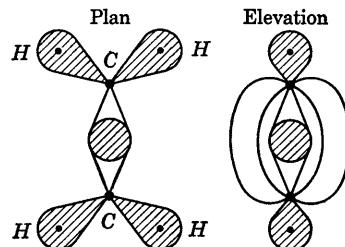


FIG. 5.17. Schematic orbitals for ethylene (C_2H_4).

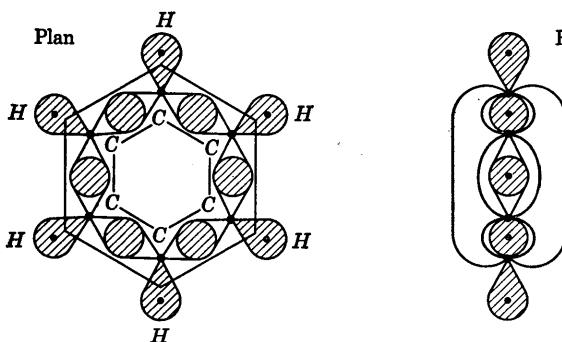


FIG. 5.18. Schematic orbitals for the benzene ring (C_6H_6).

(*b*) with one maximum along x and another along z , the cosines are proportional to $(1,0,0)$, $\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0\right)$, $\left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}, 0\right)$, $(0,0,\pm 1)$ as illustrated in Fig. 5.16. The tetragonal case is characteristic of methane, carbon tetrachloride, carbon dioxide, etc., and the trigonal case appears to apply to ethylene (Fig. 5.17), benzene (Fig. 5.18), and others. The trigonal species tend to be coplanar, and the electrons of the wave function normal to the plane combine or overlap above and below yielding additional binding energy of a type known chemically as *resonance energy*. In the schematic form of the accompanying figures it is easy to see that these bonds might be expected to have a resistance to torsion, as indeed they do. Thus the structure tends to remain planar, and, while it is possible for the two halves of the ethylene molecule to rotate about the symmetry axis relative to one another, this requires energy, and a

quantized torsional rotation is characteristic of this type of *unsaturated* double bond. The acetylene molecule, C_2H_2 , as shown in Fig. 5.19, has a triple bond between the carbon atoms, and the spectroscopic evidence indicates it is a linear molecule.

The foregoing discussion of atomic binding through the pairing of relatively free electrons with oppositely oriented spins is illustrative of what is possibly one of the simpler of molecular mechanisms. By itself it is inadequate to furnish any general description of molecular structure. There is at present no complete account of the nature of chemical binding which is in accordance with the vast amount of semiquantitative chemical data on the subject. The evolution of an adequate theory may require very fundamental developments in our knowledge of electron configurations in the presence of a plurality of ionic centers. A many-body problem of this nature is at present completely intractable. The treatment of polyatomic molecules in terms of bonds between constituent pairs is essentially approximate because of the large interdependence of these pairwise bonds on one another. However, the present quantitative approach to the problem is largely confined to analysis in terms of such bonds and the establishment of the various energy minima which such configurations may present. Even within this limited context it is necessary to discriminate between different types of electron bonds. The symmetrical bond characteristic of H_2 is probably quite different quantitatively from the very unsymmetrical assimilation of the hydrogen electron by the chlorine atom in HCl . Polyelectronic bonds are probably capable of still greater variety, and those involved in compounds of carbon, nitrogen, phosphorus, etc., may be capable of several distinct variants representing relative energy minima or metastable states of considerable stability. These bonds are likewise found to be affected both in energy and geometry by the other atoms in the molecule. A considerable empirical knowledge of these matters is represented by the rules of organic chemistry, but a quantitative unifying theory of the phenomena is at present lacking.

5.4. Molecular Nuclear Motion. The preceding section has been concerned with a simple quantitative and more general qualitative approach to the problem of the electronic motion in molecules. Approximations utilizing hydrogen atomic wave functions could be used to derive the characteristic energies $U_n(\xi)$ in terms of the separation of the atoms constituting a diatomic molecule. But this function [Eq. (5.7)] is too complicated for actual use, and simpler empirical forms which are without any direct theoretical foundation are found to yield adequate approxima-

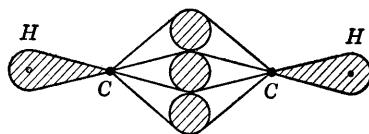


FIG. 5.19. Schematic orbitals for acetylene (C_2H_2).

tions. One such expression for the energy of two atomic constituents as a function of separation R is the following:

$$U(R) = U_0(e^{-2(R-R_0)/a} - 2e^{-(R-R_0)/a}) \quad (5.9)$$

This form, in which U_0 , R_0 , and a are parameters characteristic of the bond, was first proposed by Morse.¹ The function may be used to represent the elastic energy of separation between the two components of a diatomic molecule (see Fig. 5.11). With somewhat less precision this function may also represent the elastic energy between any two pairs of atoms in complex molecules. A complex molecule has other types of motion in addition to the simple bond extensions, for example bond bending energies and hindered or unhindered rotation energies of one subgroup of atoms with respect to another. All these possibilities present a very complex picture of the general polyatomic molecule and its quantum-mechanical description. Some progress in its theoretical analysis can be made with the help of simplifying assumptions and group theory. Nevertheless, the problem is quite difficult to any approximation beyond that in which portions of the polyatomic molecule are assumed to be simple diatomic molecules.

For the purpose of analyzing its motion, a molecule may be thought of as an extended elastic body consisting of an aggregate of mass points, representing the nuclei and their close electronic shells whose relative positions are determined by the elastic forces exerted by the ambient structure of the outer electrons. The total angular momentum consists of that of the electron structure together with that of the rotational motion of the nuclei and their close electronic shells about their common center of gravity. The discussion of the general polyatomic molecule is rather involved since it resembles a general rigid body in having three different moments of inertia. No torques are assumed, and hence the angular momentum is a constant of the motion; but the instantaneous axis about which the rotation takes place is in general constantly changing. In a frame of reference fixed in space in which the center of gravity of the molecule is at rest, the motion is described as in mechanics by the rolling without slipping of the momental ellipsoid with its center fixed, upon a fixed plane normal to the angular momentum vector. This is the general unsymmetrical top problem. If the molecule is such that it has an axis of symmetry, i.e., two of its moments are alike, the ellipsoid is one of revolution and the motion is much simpler and similar to that of the symmetrical top. Examples are diatomic and linear polyatomic molecules with values of $\Lambda > 0$, that is, with an angular momentum about the axis of symmetry. Here in the steady motion the instantaneous axis of rotation rotates uniformly about the fixed angular-momentum vector;

¹ P. M. Morse, *Phys. Rev.*, **34**, 57 (1929).

the motion is that of two rough cones, their apexes coinciding, that roll upon one another.¹ In the particular case that $\Lambda = 0$ the only rotation possible is that about an axis normal to the symmetry axis and the motion is that of a simple rotator with a fixed axis.

The subsequent discussion is limited essentially to diatomic molecules in order to present a theory in fairly simple form and illustrate its adequacy by comparison with typical experimental data. The method of handling Eq. (5.5'') depends on whether one considers a top or a simple rotator. The former is rather more involved than the latter, and in the interest of simplicity the diatomic molecule will be considered as a simple rotator; but the effect of an angular momentum about the axis of symmetry will be pointed out. Following this course, the first step is to write Eq. (5.5'') in terms of the reduced mass M_r , where $1/M_r = 1/M_1 + 1/M_2$, M_1 and M_2 being the masses of the two constituent atoms. The equation then becomes

$$\nabla^2 u(\xi) + \frac{2M_r}{\hbar^2} [W - U(R)] u(\xi) = 0 \quad (5.10)$$

The letter ξ stands for the angles ϕ and θ and the separation R , as shown in Fig. 5.20, and ∇^2 is given as for the hydrogen atom by

$$\nabla^2 = \frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial}{\partial R} \right) + \frac{1}{R^2 \sin \theta} \frac{1}{\theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{R^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}$$

In fact the problem is the same as that for hydrogen except that the electron mass occurring in that treatment is much less than the mass of a nucleus occurring here and the dependence of the potential energy on R is given by Eq. (5.9) rather than by the Coulomb law. Following exactly the procedure of Sec. 3.4, $u(\xi)$ may be written as $S(R)\Theta(\theta)\Phi(\phi)$ and the angular functions satisfy the equations

$$\begin{aligned} \frac{\partial^2 \Phi}{\partial \phi^2} + M^2 \Phi &= 0 \\ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right) + \left[J(J+1) - \frac{M^2}{\sin^2 \theta} \right] \Theta &= 0 \end{aligned}$$

¹ See for instance J. L. Synge and B. A. Griffith, "Principles of Mechanics," McGraw-Hill Book Company, Inc., New York, 1942; H. Goldstein, "Classical Mechanics," Addison-Wesley Publishing Company, Cambridge, Mass., 1950.

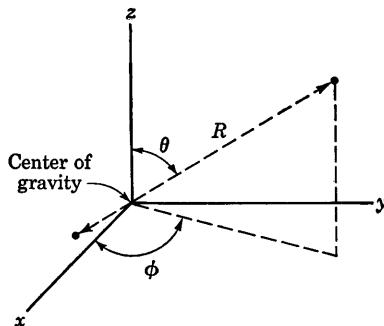


FIG. 5.20. Construction for diatomic molecules.

where M is written for m of Eq. (3.10) and J for l of Eq. (3.11) in accordance with the accepted notation for the molecular problem. The angular functions are then

$$\Phi = A_\phi e^{iM\phi}$$

and

$$\Theta = A P_J^{|M|}(\cos \theta)$$

For normalization the constants are the following:

$$A_\phi = \frac{1}{\sqrt{2\pi}} \quad \text{and} \quad A_\theta = \left[\frac{(2J+1)(J-|M|)!}{2(J+|M|)!} \right]^{\frac{1}{2}}$$

M takes on integral values and so does J , where $J \geq |M|$. The quantity $J(J+1)\hbar^2$ is the square of the total angular momentum which is a constant of the motion and appears in the equation for $S(R)$:

$$\frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial S}{\partial R} \right) + \frac{2M_r}{\hbar^2} \left[W - U(R) - \frac{J(J+1)\hbar^2}{2M_r R^2} \right] S = 0 \quad (5.11)$$

It might be pointed out here that if there is an electronic angular-momentum component along the axis of symmetry the above equation is modified by an additional term and the following classical consideration yields the correct result. In Sec. 3.4 it was seen that the term $J(J+1)\hbar^2/2M_r R^2$, the square of the angular momentum over twice the moment of inertia, represents the energy associated with classical rotation. In terms of the components of angular momentum about the three principal axes having moments of inertia I_x , I_y , and I_z , the classical energy of rotation is

$$T = \frac{1}{2} \left(\frac{L_x^2}{I_x} + \frac{L_y^2}{I_y} + \frac{L_z^2}{I_z} \right)$$

where L_x , L_y , and L_z are the components of angular momentum about these three axes. The total angular momentum squared,

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

is constant under no torque and is to be identified with $J(J+1)\hbar^2$. The square of the component about the axis of symmetry, say the z axis, arising from the electronic motion has been written Λ^2 or $\lambda^2\hbar^2$ which is to be identified with L_z^2 if there is no nuclear rotation about this axis. For a diatomic or linear molecule there is no such rotation since the nuclei form merely a line of point charges. Also for such a molecule the symmetry assures that $I_x = I_y = I$, and since

$$L^2 = J(J+1)\hbar^2$$

$$L_x^2 + L_y^2 = J(J+1)\hbar^2 - \Lambda^2$$

so

$$T = \frac{1}{2} \left[\frac{J(J+1)\hbar^2}{I} + \Lambda^2 \left(\frac{1}{I_z} - \frac{1}{I} \right) \right]$$

I_z is the effective electronic moment of inertia about z , and for a diatomic molecule $I = M_r R^2$. Thus classically one would expect the rotational energy expression to contain the additional term in Λ , and this is borne out by a quantum-mechanical analysis of the symmetrical top problem. Therefore the wave equation for the nuclear motion of a diatomic molecule including the possibility of electronic orbital angular momentum is

$$\frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial S}{\partial R} \right) + \frac{2M_r}{\hbar^2} \left[W - U(R) - \frac{J(J+1)\hbar^2}{2M_r R^2} + \frac{\Lambda^2}{2} \left(\frac{1}{M_r R^2} - \frac{1}{I_z} \right) \right] S = 0 \quad (5.11')$$

It is evident that if Λ is not equal to zero the possible angular-momentum values for the molecule do not run up from $J = 0$ but up from that value of J for which $J(J+1)\hbar^2 > \Lambda^2$. This means that those nuclear rotational energy levels for which this inequality does not hold do not exist, but otherwise the presence of Λ does not modify the energy structure appreciably. The existence of Λ does have one other consequence through its effect upon the dipole-radiation transition rules. The discussion of the angular term above is closely parallel to that of the hydrogen atom in Sec. 3.4 so that if the molecule is polar, i.e., there is a separation of the centers of positive and negative charge, dipole radiation emission would be expected for a change in J by one unit, that is, $\Delta J = \pm 1$. However, if Λ can also change, this essentially means that there can be a change in angular momentum of the dipole and hence a radiation of energy even if J does not change, i.e., the selection rules in the presence of Λ are $\Delta J = 0, \pm 1$ but, if $\Lambda = 0$, $\Delta J = \pm 1$ only.

Except for the above points the existence of Λ is without effect, and it will be taken as zero for simplicity in the following. To complete the analysis the problem is to solve Eq. (5.11) with some acceptable form of $U(R)$ such as that given by Eq. (5.9) or some satisfactory approximation to it. The form of Eq. (5.9) is unnecessarily complicated for small amplitudes of vibration, and it can be expanded in powers of $(R - R_0)/a$ to yield

$$\begin{aligned} U(\rho) &= U_0 \left(1 - \frac{2\rho}{a} + \frac{4\rho^2}{2a^2} - \frac{8\rho^3}{6a^3} \dots \right) - 2 \left(1 - \frac{\rho}{a} + \frac{\rho^2}{2a^2} - \frac{\rho^3}{6a^3} \dots \right) \\ &= -U_0 \left(1 - \frac{\rho^2}{a^2} + \frac{\rho^3}{a^3} + \dots \right) \end{aligned}$$

to terms of this order, where $\rho = R - R_0$. This will be taken as a satisfactory approximation for the present discussion. Also in terms of ρ

$$\frac{1}{R^2} \cong \frac{1}{R_0^2} \left(1 - \frac{2\rho}{R_0} + \frac{3\rho^2}{R_0^2} \right)$$

to this approximation. Finally if $X(R) = RS(R)$,

$$\frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial S}{\partial R} \right) = \frac{1}{R} \frac{\partial^2 X}{\partial R^2}$$

and Eq. (5.11) becomes

$$\frac{\partial^2 X(\rho)}{\partial \rho^2} + \frac{2M_r}{\hbar^2} \left[\underline{\underline{W + U_0}} - \frac{\underline{\underline{U_0\rho^2}}}{\underline{\underline{a^2}}} + \frac{\underline{\underline{U_0\rho^3}}}{\underline{\underline{a^3}}} - \frac{J(J+1)\hbar^2}{2M_r R_0^2} \left(\underline{\underline{1}} - \frac{\underline{\underline{2\rho}}}{\underline{\underline{R_0}}} + \frac{\underline{\underline{3\rho^2}}}{\underline{\underline{R_0^2}}} \right) \right] X = 0 \quad (5.12)$$

From Table 5.1, a is seen to be of the same order of magnitude as R_0 for most molecules. Also the ratio of $J(J+1)\hbar^2/2M_r R_0^2$ to U_0 is of the order of the electron-nuclear mass ratio. Thus the triply underlined term is largest, the doubly underlined terms are next in importance, and the singly underlined terms are higher-order corrections. Neglecting these last correction terms and writing

$$x^2 = \frac{\sqrt{2M_r U_0}}{\hbar a} \rho^2 \quad \text{and} \quad \sigma = \frac{a}{\hbar} \sqrt{\frac{2M_r}{U_0}} \left[W + U_0 - \frac{J(J+1)\hbar^2}{2M_r R_0^2} \right]$$

the equation becomes

$$\frac{\partial^2 X}{\partial x^2} + (\sigma - x^2)X = 0 \quad (5.12')$$

to this approximation. This equation corresponds to the harmonic-oscillator approximation for which the energy $U(\rho)$ is considered to be represented adequately by a quadratic function of ρ . The angular frequency ω_c associated with a classical harmonic oscillator having a potential energy of the form $U_0\rho^2/a^2$ is $\frac{1}{a} \sqrt{\frac{2U_0}{M_r}}$. In terms of ω_c the symbols previously used become

$$x = \frac{\rho}{\rho_0}, \quad \rho_0 = \sqrt{\frac{\hbar}{\omega_c M_r}}, \quad \omega_c = \frac{1}{a} \sqrt{\frac{2U_0}{M_r}} \\ W = -U_0 + \frac{\sigma \hbar \omega_c}{2} + \frac{J(J+1)\hbar^2}{2M_r R_0^2}$$

Equation (5.12') is of a standard form and can be solved conveniently by the substitution $X = CH(x)e^{-x^2/2}$, where C is an arbitrary constant, yielding

$$\frac{\partial^2 H}{\partial x^2} - 2x \frac{\partial H}{\partial x} + (\sigma - 1)H = 0$$

Assuming a series representation of H ,

$$H = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$$

and applying the condition that the coefficient of each power of x in the resulting equation must vanish, the following condition is found for the a 's:

$$\frac{a_{n+2}}{a_n} = \frac{2n + 1 - \sigma}{(n + 2)(n + 1)} \quad (5.13)$$

If the series for H does not terminate, it can be shown that X becomes infinite as x approaches infinity so the numerator of Eq. (5.13) must vanish for some n , that is, $\sigma = 2n + 1$ if the wave functions are to be satisfactory. Thus it is seen that the finiteness of the wave function determining the nuclear positions in the molecule restricts the values of the energy to

$$W_{nJ} = -U_0 + (n + \frac{1}{2})\hbar\omega_c + \frac{J(J + 1)\hbar^2}{2M_r R_0^2} \quad (5.14)$$

where n is an integer. This is analogous to the quantum conditions for E_n derived in Sec. 3.4 for the hydrogenlike atom. By an argument analogous to that of Sec. 4.6, the solutions of Eq. (5.12') are seen to be orthonormal, and the coefficients may be found from $\int_{-\infty}^{\infty} X_n^2(x) dx = 1$. In the harmonic-oscillator approximation the domain of x is from $-\infty$ to $+\infty$, and in terms of the coefficients so determined

$$S_n(R) = \frac{X}{R} = \frac{1}{R} (\sqrt{\pi} \rho_0 2^n n!)^{-\frac{1}{2}} H_n \left(\frac{\rho}{\rho_0} \right) e^{-(\rho/\rho_0)^2/2} \quad (5.15)$$

where H_n is the polynomial given by the recursion formula [Eq. (5.13)] in which $a_0 = 1$ and $a_1 = 2$ terminating with the term x^n . Of course only even or only odd terms occur in this polynomial, not both, so H_n and hence S_n are even or odd in parity. By writing out a few terms, an alternative equivalent definition of $H_n(x)$ may be seen to be

$$H_n(x) = (-1)^n e^{x^2} \frac{\partial^n}{\partial x^n} e^{-x^2}$$

5.5. Molecular Spectroscopy. Diatomic molecular energy levels are given to a first approximation by Eq. (5.14). The first term of this

equation is the energy of the electronic structure as measured from a zero of the separate atoms, the second is associated with the nuclear vibratory motion, and the third and smallest term with the rotation of the molecule. The vibrational levels are equally spaced, the difference

$$W_{(n+1),J} - W_{n,J} = \hbar\omega_c$$

being simply the classical angular frequency associated with the assumed potential-energy curve times the constant \hbar . The rotational levels, on the other hand, show an increasing spacing, i.e.,

$$W_{n,J+1} - W_{n,J} = \frac{2(J + 1)\hbar^2}{2M_r R_0^2}$$

The lowest energy level for a particular U_0 , which is that for which $n = J = 0$, is $-U_0 + \frac{1}{2}\hbar\omega_c$ to this approximation. Thus there is a slight vibratory motion, known as the *zero-point oscillation*, in the lowest or zero-point vibrational state. The classical amplitude of oscillation for this energy is obtained by setting the energy $\frac{1}{2}\hbar\omega_c$ equal to the potential energy at maximum extension, say ρ_0 , which is $U_0\rho_0^2/a^2$. This yields $\rho_0 = (\hbar/M_r\omega_c)^{\frac{1}{2}}$. Thus ρ_0 corresponds with the earlier use of this symbol, and from the expression for $S(R)$ of the previous section ρ_0 is seen to be the distance at which $S(R)$ drops to $e^{-\frac{1}{2}}$ of its maximum value for the lowest energy state.

The selection rules determining the electric dipole transitions that can take place between electronic levels have been mentioned and also those that determine transitions between rotational levels, $\Delta J = 0, \pm 1$. There are similar rules that determine dipole-radiation transitions between vibrational levels that can be derived following a classical analogy on the assumption that the radiation amplitude is proportional to the dipole moment P associated with the wave functions of the initial and final vibrational states (Sec. 3.6). Thus the dipole-radiation amplitude is proportional to

$$\int X_n P X_m dR$$

If the molecule is a polar molecule like HCl it has a dipole moment, say P_0 , in its normal state. Even if the molecule normally has no dipole moment, nuclear vibrations may give rise to one unless the molecule is very symmetrical, as in such cases as H_2 , N_2 , etc. A dipole moment arising from a variation of internuclear distance may be taken as proportional to this variation to a first approximation, and P may be written as $P_0 + e'\rho$. The dipole-radiation amplitude would then be proportional to

$$(\pi\rho_0^2 2^n 2^m n! m!)^{-\frac{1}{2}} \int (P_0 + e'\rho) H_n(x) H_m(x) e^{-x^2} d\rho$$

or

$$(\pi 2^n 2^m n! m!)^{-\frac{1}{2}} (P_0 \int H_n H_m e^{-x^2} dx + e' \rho_0 \int x H_n H_m e^{-x^2} dx)$$

The first integral is unity if $n = m$ and vanishes otherwise. This term represents no change in the vibrational motion and is associated, for instance, with the pure rotation spectrum of a polar molecule. The second term may be evaluated by differentiating

$$H_n(x) = (-1)^n e^{x^2} \frac{\partial^n}{\partial x^n} e^{-x^2}$$

which yields $\partial H_n / \partial x = 2xH_n - H_{n+1}$. Differentiating again and using the differential equation for H with $\sigma = 2n + 1$, it is seen that

$$2xH_n(x) = 2nH_{n-1}(x) + H_{n+1}(x)$$

Using this expression for the product of x and one H in the second integral and recalling that the X 's are normalized, it is seen that there are the following nonvanishing terms in the expression for the dipole-radiation amplitude:

$$P_0, \quad + \frac{e'\rho_0}{\sqrt{2}} (n+1)^{\frac{1}{2}}, \quad + \frac{e'\rho_0}{\sqrt{2}} n^{\frac{1}{2}}$$

The three terms arise from $m = n$, $m = n + 1$, and $m = n - 1$ respectively. Thus in the harmonic-oscillator approximation, if a dipole moment is induced by a change in the nuclear separations, the selection rule for dipole radiation resulting from a change in vibrational motion is $\Delta n = \pm 1$. The intensities for the transitions $\Delta n = +1$ and $\Delta n = -1$ are in the ratio $(n+1)/n$. It should be emphasized, however, that Eq. (5.12'), upon which the foregoing discussion rests, is only a first approximation to that for the actual problem and the vibratory motion of the nuclei takes place under the influence of a potential function which is not of the symmetric parabolic form assumed in approximating Eq. (5.12) by Eq. (5.12'). Thus the selection rule is not to be regarded as rigorous. Also $e'\rho$ vanishes for the free vibration of a highly symmetrical molecule made up of like atoms such as H_2 or O_2 , and no dipole radiation is emitted.

Infrared Spectra. Those transitions in which the electronic structure is unaltered and the change in energy is limited to a change in vibrational amplitude or rotational angular momentum represent radiation in the long-wavelength, or infrared, region of the spectrum. By analogy with classical dipole-radiation theory, radiation would be expected only if a relative displacement of positive and negative charge is associated with such motion. Thus symmetrical molecules such as H_2 , O_2 , etc., should not generate such spectra in the absence of external forces since neither elongation nor rotation induces a dipole moment, and no infrared spectra are observed from such molecules. On the other hand, an unsymmetrical

molecule such as HCl has a dipole moment and both vibrational and rotational spectra are observed. For instance, Czerny has observed strong absorption at the wave numbers given in Table 5.2 for HCl in the

TABLE 5.2. FAR-INFRARED HCl ABSORPTION BANDS

<i>S</i>	$\bar{\nu}_{\text{obs}}, \text{cm}^{-1}$	$\Delta\bar{\nu}, \text{cm}^{-1}$	$S \Delta\bar{\nu}_{\text{av}}, \text{cm}^{-1}$
1			20.47
2			40.94
3			61.41
4	83.03		81.88
5		2×20.63	102.37
6	124.30	20.73	122.82
7	145.03	20.48	143.29
8	165.51	20.36	163.76
9	185.86	20.52	184.23
10	206.38	20.12	204.75
11	226.50		225.17

far infrared. The first differences are seen to be nearly constant, suggesting the serial numbers at the left, and it is seen that $S \Delta\bar{\nu}_{\text{av}}$ agrees fairly well with the observations. This is in accord with the hypothesis that the absorption bands are due to transitions between rotational energy levels where the wave numbers should be given theoretically by

$$\bar{\nu}_{\text{obs}} = \bar{\nu}_{J+1} - \bar{\nu}_J = \frac{1}{hc} (W_{J+1} - W_J) = \frac{(J+1)\hbar}{2\pi c M_r R_0^2}$$

The difference between wave numbers in the absorption spectrum should then be $\hbar(2\pi c M_r R_0^2)^{-1}$. If this is equated to the best present value of $\Delta\bar{\nu} = 21.182$, the moment of inertia is found to be

$$M_r R_0^2 = 2.642 \times 10^{-47} \text{ kg-m}^2$$

Using the masses of H^1 and Cl^{35} , $M_r = 1.61 \times 10^{-27} \text{ kg}$ so

$$R_0 = 1.28 \times 10^{-10} \text{ m}$$

This is the correct order of magnitude confirming the identification of these bands and the general theory of molecular rotation spectra. The frequency is very sensitive to the mass of the lighter atom, and as M_r is nearly twice as great for deuterium chloride the absorption bands of D^2Cl^{35} are found at about twice the wavelength of H^1Cl^{35} . With the application of modern radio-frequency techniques, measurements of this type have been pushed up into the centimeter range and interesting information has been gained about more complicated molecules. However, as will be seen later, the low rotational levels of heavy diatomic

molecules are so sparsely populated that their absorption of radiation is difficult to observe. Nevertheless, Townes and his collaborators¹ have measured the frequencies corresponding to transitions from the first to the second rotational states for the vaporized chlorides of Na, K, Cs, and Tl. For the lowest vibrational states of NaCl³⁵ and NaCl³⁷, for instance, these frequencies are found to be 26,051.1 and 25,493.9 Mc, respectively. The ratio of these frequencies is equal to the ratio of the reduced masses in accordance with the preceding formula, and the nuclear separation R_0 is found to be 2.36×10^{-10} m.

In the near-infrared region there are also HCl absorption bands, the one at longest wavelength consisting of a group of some 25 lines lying between 3.2 and 3.9 μ . From Eq. (5.14) the anticipated wave numbers of rotation-vibration lines would be

$$\bar{\nu}_{\Delta J=1} = \bar{\nu}_{n+1,J-1} - \bar{\nu}_{n,J} = \frac{\omega_c}{2\pi c} - \frac{J\hbar}{2\pi c M_r R_0^2} \quad \text{where } J = 1, 2, 3, \dots$$

$$\bar{\nu}_{\Delta J=0} = \bar{\nu}_{n+1,J} - \bar{\nu}_{n,J} = \frac{\omega_c}{2\pi c}$$

$$\bar{\nu}_{\Delta J=-1} = \bar{\nu}_{n+1,J+1} - \bar{\nu}_{n,J} = \frac{\omega_c}{2\pi c} + \frac{(J+1)\hbar}{2\pi c M_r R_0^2} \quad \text{where } J = 0, 1, 2, \dots$$

These of course neglect any possible change in R_0 due to a finite amplitude of anharmonic vibration or to the centrifugal force of molecular rotation. Since J refers to the lower state, $\Delta J = 1$ is not possible for $J = 0$. The line at $\omega_c/2\pi c$ is missing if the transition $\Delta J = 0$ is not permitted. The wave numbers of the 12 central lines in the band of about 25 lines for the HCl band mentioned above are given in Table 5.3.

It is seen that $\Delta\bar{\nu}$ is not constant but shows a trend. The next approximation in the theory would have to be used to account for this trend. Nevertheless, the average value of $\Delta\bar{\nu}$ is closely the same as in the previous table, which identifies it with rotational energy. The serial number $S = 0$ is missing indicating that $\Delta J = 0$ is not permitted. From this, it may be inferred that there is no electronic angular momentum since $\Delta J = 0$ is permitted only if Λ is not zero. This is analogous to the situation in atomic spectra in which $\Delta J = 0$ is permitted if the total angular momentum is made up of both spin and orbital types. Dipole radiation due to a change in orbital angular momentum may take place with no change in J if a compensating change in spin angular momentum occurs. In molecules the molecular rotation may change with accompanying dipole radiation with no change in J if the electronic angular momentum also changes by an equal and opposite amount during the transition.

¹ M. L. Stitch, H. Honig, and C. H. Townes, *Phys. Rev.*, **86**, 813 (1952).

There are additional HCl bands of the same nature nearer the visible, indicating higher vibrational-energy changes than $\Delta n = 1$. The selection rule limiting Δn to this value depends upon the approximation here used in which the dependence of energy upon higher powers of ρ/a than

TABLE 5.3. CENTRAL COMPONENTS OF A NEAR INFRARED HCl ABSORPTION BAND (*after Herzberg*)

S	$\bar{\nu}_{\text{obs}}, \text{cm}^{-1}$	$\Delta \bar{\nu}_{\text{obs}}, \text{cm}^{-1}$	$\Delta(\Delta \bar{\nu}_{\text{obs}}), \text{cm}^{-1}$
6	2997.78		
5	80.90	16.88	
4	63.24	17.66	0.78
3	44.89	18.35	0.69
2	25.78	19.11	0.76
1	06.25	19.53	0.42
0			
-1	2865.09	(41.16)	
-2	43.56	21.53	
-3	21.49	22.07	0.54
-4	2798.78	22.71	0.64
-5	75.79	22.99	0.28
-6	52.03	23.76	0.77

$$(S=0) \quad \bar{\nu}_{\text{calc}} = 2885.9 \text{ cm}^{-1}, \quad (\Delta \bar{\nu}_{\text{obs}})_{\text{av}} = 20.48 \text{ cm}^{-1}, \quad \Delta(\Delta \bar{\nu}_{\text{obs}})_{\text{av}} = 0.61 \text{ cm}^{-1}$$

the second is neglected. If these terms representing anharmonicity are retained, Δn can take larger values as shown below:

$\Delta n = 1$	$\bar{\nu}_{\text{obs}}, \text{cm}^{-1}$	$\Delta \bar{\nu}_{\text{obs}}, \text{cm}^{-1}$	$\Delta(\Delta \bar{\nu}_{\text{obs}}), \text{cm}^{-1}$
2	2885.9		
2	5668.0	2782.1	
3	8347.0	2679.0	-103.1
4	10922.7	2575.7	-103.3

The fact that $\Delta \bar{\nu}$ becomes smaller for higher values of Δn shows that the actual motion is anharmonic confirming the explanation that these higher values of Δn exist at all. The addition of a quadratic term in n to Eq. (5.14) would evidently yield a constant second difference as is observed. Thus, neglecting the term in J in that equation, a better empirical expression for W_n is

$$W_n = -U_0 + \hbar\omega_c[(n + \frac{1}{2}) + \alpha(n + \frac{1}{2})^2]$$

where α is a constant, small compared to unity. The observed wave numbers for a transition Δn to the $n = 0$ state are

$$\bar{\nu}_{\text{obs}} = \frac{W_{n+\Delta n} - W_n}{hc} = \bar{\nu}_v[\Delta n + \alpha \Delta n^2 + \alpha \Delta n]$$

where $\bar{\nu}_v = \frac{\omega_c}{2\pi c}$. The first differences are

$$\Delta\bar{\nu}_{\text{obs}} = \bar{\nu}_v(1 + 2\alpha \Delta n + 2\alpha)$$

and the second differences are

$$\Delta(\Delta\bar{\nu}_{\text{obs}}) = \bar{\nu}_v 2\alpha$$

which are seen to be independent of Δn in accord with the table. $\bar{\nu}_v$ is simply related to $\bar{\nu}_{\text{obs}}(\Delta n = 1)$ by

$$\bar{\nu}_{\text{obs}}(\Delta n = 1) = \bar{\nu}_v(1 + 2\alpha) = \bar{\nu}_v + \Delta(\Delta\bar{\nu}_{\text{obs}})$$

Thus in the case of HCl

$$\bar{\nu}_v = 2885.9 + 103.2 = 2989.1 \text{ cm}^{-1}$$

From the value of $\bar{\nu}_v$ or ω_c , the zero-point vibrational energy and the elastic constant for HCl can be calculated:

$$\omega_c = 2\pi c 298,910 = 5.63 \times 10^{14} \text{ rad sec}^{-1}$$

$$E_0 = \hbar\omega_c \left(\frac{1}{2} + \frac{\alpha}{4} \right) = 2.992 \times 10^{-20} \text{ joules}$$

$$= 0.186 \text{ ev}$$

$$k_{\text{elastic}} = \frac{2U_0}{a^2} = M_r \omega_c^2 = 510 \text{ kg sec}^{-2}$$

The elastic constant depends on the electron structure and therefore is the same for the different isotopes of an element; thus the characteristic vibrational frequency ω_c is inversely proportional to $\sqrt{M_r}$. The case of HCl may be used for illustration. By concentration of the isotopes, HCl can be obtained in any of four forms and four sets of vibrational bands due to these different isotopic combinations can be observed. The fractional displacements relative to H^1Cl^{35} are given below:

$$1 - \frac{\omega_c(\text{H}^1\text{Cl}^{37})}{\omega_c(\text{H}^1\text{Cl}^{35})} = 1 - \sqrt{\frac{M_r(\text{H}^1\text{Cl}^{35})}{M_r(\text{H}^1\text{Cl}^{37})}} = 0.75 \times 10^{-3}$$

$$1 - \frac{\omega_c(\text{D}^2\text{Cl}^{35})}{\omega_c(\text{H}^1\text{Cl}^{35})} = 0.283$$

$$1 - \frac{\omega_c(\text{D}^2\text{Cl}^{37})}{\omega_c(\text{H}^1\text{Cl}^{35})} = 0.284$$

The Cl^{37} shift is about 2 wave numbers per centimeter, which produces very close doublets in the rotational structure of the band. The D^2 shift is about 780 wave numbers per centimeter to a different region of the spectrum. Thus molecular constants can be calculated from the experimental data from absorption by polar molecules in the near and far infrared. They are in general agreement with the values that would be

anticipated from consistency with other experiments and provide strong confirmation of the general correctness of the theory.

Raman Spectra. Additional information on rotational and vibrational spectra comes from an analysis of the inelastic scattering of monochromatic light by molecules. An inelastic scattering process is one in which there is a net exchange of energy between the radiation and the scattering system. It is clear from this and previous chapters that if there is no ionization or dissociation (which processes are not here considered) the atomic or molecular scatterer can contribute or absorb only such an amount of energy as is represented by one of the possible differences between the initial state and some other characteristic quantized state. Thus

$$\hbar(\omega - \omega') = -(W_i - W_f)$$

where ω and ω' are the incident and scattered angular frequencies, respectively, and W_i and W_f are the initial and final energy states of the scattering system. This type of scattering with a change in wavelength or frequency was predicted by Smekal¹ and observed first by Raman² in the visible region using molecular scatterers. In principle it would be expected to be evident in X-ray scattering from atomic systems, but it has not actually been observed in such scattering.

The process taking place can be understood in somewhat more detail by adopting the classical analogy of radiation by a dipole excited by radiation of wavelength long compared with the dipole's dimensions, as in Sec. 5.1. There the first-order change in energy at such a scattering process was seen to be zero. However, the charge density times the displacement associated with a transition from a state n to a state m is not zero in general as can be seen by writing out the dipole moment developed under the influence of the radiation field. Following the nomenclature of Sec. 5.1,

$$u_n'' = u_n + \sum_k \frac{W'_{nk}}{E_n - E_k} u_k$$

$$u_m'' = u_m + \sum_j \frac{W'_{mj}}{E_m - E_j} u_j$$

where

$$W'_{nk} = \int u_k^*(eExx) u_n d\tau = eEx \int u_k^* x u_n d\tau = eEx X'_{nk}$$

Then the polarization is

$$P_{nm} = \int ex u_n'' u_m^{*''} d\tau$$

$$= eX'_{nm} + e \int \sum_k u_m^* \frac{xW'_{nk}}{E_n - E_k} u_k d\tau + e \int \sum_j u_n \frac{xW'_{mj}}{E_m - E_j} u_j^* d\tau$$

¹ A. Smekal, *Naturwissenschaften*, **11**, 873 (1923).

² C. V. Raman, *Indian J. Phys.*, **2**, 387 (1928).

The last two terms are of the same form, and P_{nm} may be written

$$P_{nm} = eX'_{nm} + E_x e^2 \sum_k \left(\frac{X'_{nk} X'_{mk}}{E_n - E_k} + \frac{X'_{nk} X'_{mk}}{E_m - E_k} \right)$$

The first term, which is of the first order in e , represents induced absorption or emission of the frequency $\nu_{nm} = (E_n - E_m)/\hbar$ under the influence of the radiation field. The latter two terms correspond to processes in which the system changes from the state n to the state m by way of the intermediate state k . Thus it is a two-step or *two-quantum* type of process. If $n = m$ the terms represent a return to the initial state as a result of absorbing and reemitting the same frequency. This is called *resonance scattering* and it involves no net exchange of energy between the radiation and excitation forms. If $n \neq m$ the initial and final states are different and the change in excitation energy of the molecular system is balanced by the change in the radiant energy before and after the process as indicated in Fig. 5.21. Thus the incident energy is increased or decreased by an amount $h\nu_{nm}$ and reradiated. Such a process is known as *Raman scattering*.

The occurrence of a Raman transition depends on the nonvanishing of at least a pair of X'' 's, one involving the initial and the other the final state with an intermediate state in common as was mentioned in Sec. 3.6. It is clear from the form of X' that as the coordinate x is antisymmetric the integral for X' will vanish or not depending on whether the product $u_k^* u_n$ is symmetric or antisymmetric in the coordinate. Each u may be thought of as the product of an electronic function u_e , which is a solution of Eq. (5.5'), and a nuclear function, written as the product of a rotational and vibrational function $u_r u_v$, which is a solution of Eq. (5.5''). Thus if the product $(u_e u_r u_v)_k^* (u_e u_r u_v)_n$ is symmetric the transition is forbidden, and if it is antisymmetric the transition may occur. u_v is a function of the separation of the nuclei only, and hence it is without influence upon the relative sign of the contribution to the integral X'_{nk} made by volume elements oppositely situated with respect to the origin of the angular variables. The magnitude of the product $u_r u_e$ will be equal for points oppositely situated with respect to the origin, but the sign may be either the same or opposite. As x is odd with respect to reflection in the origin, X'_{nk} is finite only if $(u_e u_r)_n$ is of the opposite symmetry to $(u_e u_r)_k$. Similarly, $(u_e u_r)_m$ is of opposite symmetry to $(u_e u_r)_k$, and in accordance with the selection rule for angular variables the angular momentum increases

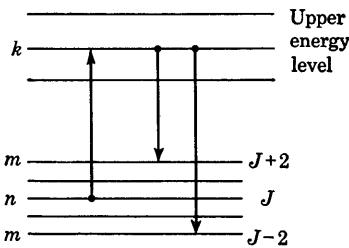


FIG. 5.21. Schematic of Raman transition.

or decreases by one unit at each individual process. If at the conclusion of the two processes the molecule returns to the same electronic state, $u_{en} = u_{em}$, and in consequence u_{rn} must be of the same symmetry as u_{rm} . The angular momentum in a Raman process is unchanged or changes by two units; thus $\Delta J_{nm} = 0, \pm 2$. This is an example of the rule of mutual exclusion which states that the rotational transitions that are allowed in the infrared spectrum are excluded from the Raman-scattered spectrum and vice versa.

The relation of the function u_v to the integral X'_{nk} is evidently of a somewhat different nature. The symmetry arguments applicable to the angular variables are not relevant to the vibrational motion, but the nature of the polarization associated with two vibrational states was discussed in Sec. 5.5. The relative vibrational displacement $x = \rho/\rho_0$ of that section is not to be confused with the coordinate x of the immediately preceding paragraphs. From Sec. 5.5 it is seen that the contribution to the integral for X'_{nk} of vibrational states of like quantum number is finite, and, if an electric dipole moment develops as a result of the relative displacement of the nuclei, neighboring vibrational states make finite contributions as well. Thus, though these selection rules are not as rigorous as the rotational ones owing to the anharmonicity of the

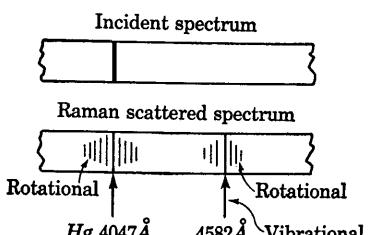


FIG. 5.22. Raman scattering from HCl.

vibrational motion, it would be expected that the intensities of the vibrational contributions to the products $X'_{nk}X'_{km}$ would exhibit decreasing magnitudes for changes of 0, 1, 2, etc., units in the vibrational quantum number. This is in accordance with the experimental results. In fact, changes in the vibrational quantum number of two or more units corre-

spond to such low intensities that they have not been observed. The general requirement for the occurrence of the Raman effect in either rotation or vibration is evidently that a molecular dipole moment should be developed under the influence of the exciting radiation; the existence of a permanent dipole moment is unnecessary. Thus the Raman effect is observed in such molecules as O₂ and N₂ which emit no infrared spectra.

The general appearance of the spectrum of light from a mercury arc scattered from HCl in the near ultraviolet is indicated in Fig. 5.22. There is one strong line in the scattered spectrum not present in the original spectrum. The wave-number difference is

$$\frac{10^8}{4047} - \frac{10^8}{4582} = 2886 \text{ cm}^{-1}$$

which agrees with the fundamental HCl vibration frequency as observed in the infrared. Wood and Dieke¹ also observed rotational Raman structure near the 4046-A line as shown in Table 5.4. This agrees with the rotational Raman theory, for

$$\Delta\bar{\nu}_{(J+2) \rightarrow J} = \frac{(J+2)(J+3)}{2} \bar{\nu}_r - \frac{J(J+1)}{2} \bar{\nu}_r = (2J+3)\bar{\nu}_r$$

where from the earlier discussion $\bar{\nu}_r = 21 \text{ cm}^{-1}$. The serial number S represents $J + \frac{3}{2}$. Only one vibrational Raman line is observed and that at a longer wavelength (a positive serial number in the sense of the table below). This is easily understood if the HCl molecule doing the

TABLE 5.4. RAMAN SCATTERED COMPONENTS FROM HCl

S	ν_{nm}, cm^{-1}	$\Delta\nu_{nm}, \text{cm}^{-1}$
-4	232.2	48.9
-3	183.3	39.5
-2	143.8	
1	-101.1	41.6
2	-142.7	44.8
3	-187.5	41.9
4	-229.4	41.6
5	-271.0	41.9
6	-312.9	39.1
7	-352.0	

scattering is initially in the lowest vibrational state for then it can only return to that same state or to a higher one since no lower one exists. Lines of both greater and smaller wave number occur in the Raman rotational structure. The scattering atom at normal temperatures has a good chance of being in a rotational state higher than the lowest state; thus, in returning, it can go to a lower state than the initial one, liberating more energy than it received as in the cases of the negative serial numbers in the table. Scattered lines of longer wavelength representing a loss of energy to the scatterer are known as *Stokes* lines. Early experiments by Stokes on the fluorescence of materials under the influence of ultraviolet radiation indicated that the fluorescent light was always of longer wavelength or lower frequency than the exciting radiation. This is known as *Stokes's law*, and the present nomenclature arises from the analogy between the fluorescent and Raman phenomena in that the wavelength of the emergent light is different from the incident light in both cases. The lines of shorter wavelength than the incident light,

¹ R. W. Wood and G. H. Dieke, *Phys. Rev.*, **35**, 1355 (1930).

which depend for their existence on the contribution of energy from the scatterer, are known as *anti-Stokes* lines. Raman-scattering experiments are very informative in molecular structure problems particularly since they give information on the classes of molecules which are not naturally polar but become so under the influence of the electric field of the light being scattered.

Electronic Spectra. Molecular spectra in the visible and ultraviolet are due to transitions in which the electronic structure changes. However, the pattern of each band is determined by accompanying changes in the nuclear motion. From Eq. (5.14),

$$\bar{\nu} = \frac{1}{2\pi\hbar c} \left\{ -(U_{0i} - U_{0f}) + [(n_i + \frac{1}{2})\hbar\omega_{ci} - (n_f + \frac{1}{2})\hbar\omega_{cf}] + \left[\frac{J_i(J_i + 1)\hbar^2}{2M_iR_i^2} - \frac{J_f(J_f + 1)\hbar^2}{2M_fR_f^2} \right] \right\} \quad (5.16)$$

where the subscripts *i* and *f* refer to the initial and final molecular states, respectively. The first term in Eq. (5.16), $U_{0i} - U_{0f}$, is the principal one in order of magnitude, and as we have seen it is not calculable in general. There are selection rules that determine the states between which dipole transitions are permitted. These are much the same as those governing atomic transitions though they differ because of the different symmetries of the two cases. The dipole polarization integral $\int u_a x u_b^* d\tau$ vanishes if the *u*'s are of the same symmetry because *x* is odd. Thus in the subscript and superscript notation of Sec. 5.3 the selection rules are

$$\begin{array}{ccc} + \rightarrow - & + \not\rightarrow + & - \not\rightarrow - \\ u \rightarrow g & u \not\rightarrow u & g \not\rightarrow g \end{array} \quad (5.17)$$

In case the nuclei are identical, their symmetry properties also become important, but this matter is postponed till the next section. Of course the rule $\Delta J = 0, \pm 1$ holds, but because of the anharmonicity of the oscillatory motion, particularly for large amplitudes, the rule $\Delta n = \pm 1$ is not effective.

A graphical discussion following Fig. 5.23 is more enlightening than an analytical one in the case of electronic molecular spectra. Consider a molecule in a certain vibrational state n_i . The coefficients of *x* as given by Eq. (5.13) are alternately positive and negative, and the probability function is greatest near the extreme excursions of the nuclei in agreement with the fact that a classical oscillator spends a greater fraction of its time in these positions where the motion is slowest. Also, the time during which a transition leading to visible radiation takes place is short compared to the time associated with a vibration. This is evident from the uncertainty principle $\Delta E_{elect} \Delta t_{elect} = \hbar = \Delta E_{vib} \Delta t_{vib}$, and, as

$\Delta E_{elect}/\Delta E_{vib} \cong 100$, $\Delta t_{vib}/\Delta t_{elect} \cong 100$. Hence during an electronic transition there is not time for the relative positions of the nuclei to alter appreciably and the selection rule for n is essentially that one of the vibrational termini of the upper state should be at the same nuclear separation as one of the vibrational termini of the lower state. In terms of Fig. 5.23, transitions are determined by joining vibrational-state termini by vertical lines. This is known as the *Franck-Condon principle*. In a plot of vibrational-structure intensities of a band, it evidently leads

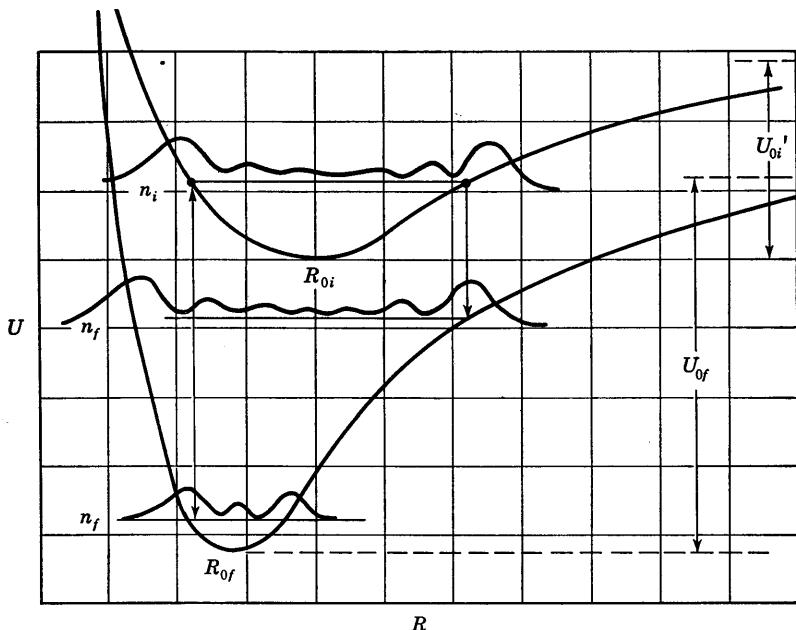


FIG. 5.23. Franck-Condon principle for electronic transitions.

to a tendency for two intense transitions from each initial state. This is in accordance with experimental observation.

In the case of absorption of radiation by a molecule in the lowest vibration state of the lowest electronic state, the maximum absorption coefficient will be for that radiation of the correct frequency to pass vertically in the diagram to the vibrational level terminating near the intersection of the vertical line and the potential-energy curve of the upper state. Thus, unless the minima of the two states are directly over one another, the molecule is generally left in a state of considerable vibrational amplitude after such an absorption process. Of course, if the absorption process takes the molecule to a point on the potential-energy curve that is above the asymptote for dissociation, the atoms have enough energy to come apart and photodissociation takes place. Obser-

vations on the transitions between the vibrational levels lead to so-called *progressions* which are the emission or absorption lines associated with one vibrational level in, say, an upper electronic state and a series of vibrational levels in the lower electronic state. From these lines the energy differences between lower vibrational levels can be derived and the limiting value of their sum, which is the dissociation energy, may be calculated. One is here concerned with the vibrational terms in the first square brackets of Eq. (5.16), but the anharmonicity necessitates the inclusion of higher-order terms in n . Thus, including a quadratic term, the wave numbers of a progression of lines ending upon successively higher vibrational levels of a final state would be

$$\bar{\nu} = \frac{1}{2\pi\hbar c} [A - (n_f + \frac{1}{2})\hbar\omega_{cf} - \alpha(n_f + \frac{1}{2})^2\hbar\omega_{cf}]$$

where A and α are constants. Then $\Delta\bar{\nu}$ for $n_f + 1$ and n_f is

$$\Delta\bar{\nu}_{(n+1),n} = \frac{\omega_{cf}}{2\pi c} [-1 - 2\alpha(n_f + 1)]$$

In the case of H_2 , 15 vibrational levels are found to exist in the lowest state and comparison with experiment yields the empirical values $\frac{\omega_{cf}}{2\pi c} = 4405.3 \text{ cm}^{-1}$ and $\alpha = -2.83 \times 10^{-2}$. Summing $\Delta\bar{\nu}$ for $n_f = 0$ to $n_f = 15$ yields

$$\begin{aligned} \sum_0^{15} \Delta\bar{\nu} &= 4405.3 \left(-15 + 2\alpha \frac{15 \cdot 16}{2} \right) \\ &= -66,097(1 - 0.448) = 36,450 \text{ cm}^{-1} \\ &= 4.52 \text{ ev} \end{aligned}$$

This is about 1 per cent above the best theoretical value and is in agreement with experiment for the dissociation energy of hydrogen. The energy interval between the bottom of the potential-energy curve and the asymptote is U_0 , which is greater than the dissociation energy by $\frac{1}{2}\hbar\omega_{cf}$; thus U_0 for the ground state of H_2 is 4.80 ev.

Methods such as the foregoing are not always unambiguous, for in many cases the potential-energy curves of different electronic states overlap or intersect. Thus for the same value of the energy and the nuclear separation (the intersection of two electronic-state curves), two electronic states can coexist and transitions from one to the other are of course likely if they are permitted by the selection rules. These will cause a breaking off of the series of vibrational levels if the intersection lies below one or both of the asymptotes before the full series scheme of

vibrational levels for the electronic state has developed. The curves of Fig. 5.24 indicate the typical possibilities. In case A the vibrational states of curve *a* would be expected to be normal below *x* but in the neighborhood of that point they should drop rapidly in intensity because of transitions to curve *b* and subsequent dissociation since *x* is above *d*. In case B where *x* is below *d* the vibrational levels are normal up to *x* and then disturbed or *perturbed* between *x* and *d*, for in this region transitions between the two electronic states can take place. Dissociation then occurs for both the levels of *a* and *b* at the energy *d*. These types of effects clearly preclude the determination of U_0 by a simple extrapolation formula.

A final point in connection with the rotational structure of electronic spectra is that the difference between the mean nuclear separation for two different electronic states gives the bands a different appearance than for the rotational structure of the vibrational bands of a single electronic state. For a certain set of values of U_{0i} , U_{0f} , n_i , and n_f , Eq. (5.16) leads to the three sets or groups of transitions for the rotational transitions $J - 1 \rightarrow J$, $J \rightarrow J$, $J + 1 \rightarrow J$ which are designated as the *P*, *Q*, and *R* branches of the band, respectively.

$$\begin{aligned}\Delta\bar{\nu}_P &= \frac{\hbar}{4\pi c M_r} \left[\frac{(J-1)J}{R_i^2} - \frac{J(J+1)}{R_f^2} \right] \\ &= \frac{\hbar}{4\pi c M_r} \left[\left(\frac{1}{R_i^2} - \frac{1}{R_f^2} \right) J^2 - \left(\frac{1}{R_i^2} + \frac{1}{R_f^2} \right) J \right] \\ \Delta\bar{\nu}_Q &= \frac{\hbar}{4\pi c M_r} \left[\frac{J(J+1)}{R_i^2} - \frac{J(J+1)}{R_f^2} \right] \\ &= \frac{\hbar}{4\pi c M_r} (J^2 + J) \left(\frac{1}{R_i^2} - \frac{1}{R_f^2} \right) \\ \Delta\bar{\nu}_R &= \frac{\hbar}{4\pi c M_r} \left[\frac{(J+1)(J+2)}{R_i^2} - \frac{J(J+1)}{R_f^2} \right] \\ &= \frac{\hbar}{4\pi c M_r} \left[\left(\frac{1}{R_i^2} - \frac{1}{R_f^2} \right) J^2 + \left(\frac{3}{R_i^2} - \frac{1}{R_f^2} \right) J + \frac{2}{R_i^2} \right]\end{aligned}$$

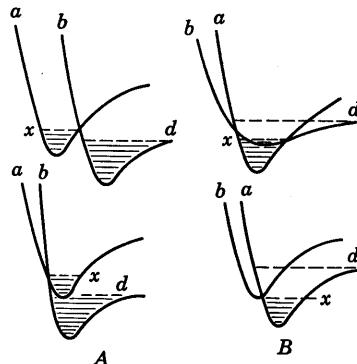


FIG. 5.24. Representative intersecting molecular-potential-energy curves.

As an illustration the wave numbers for the first nine rotational components of the CO band representing the transition ${}^1\Sigma^+ \rightarrow {}^1\Pi$ with $n_i = 0$ and $n_f = 3$ are given graphically in Fig. 5.25. (This is one part of the so-called Angstrom bands of CO.) Since there is electronic

angular momentum in the II state ($\Lambda \neq 0$), the level $J = 0$ does not exist in this state and in consequence the bands for $J = 0$ are absent. On the other hand, the transition $\Delta J = 0$ does occur (the Q branch). The $\Delta\nu$'s are quadratic functions of the index J and $\Delta\nu_R$ for $J = (J - 1)$ is the same equation as $\Delta\nu_P$ for $J = -J$. This is indicated for the experimental points by the continuation of the parabola in Fig. 5.25. A

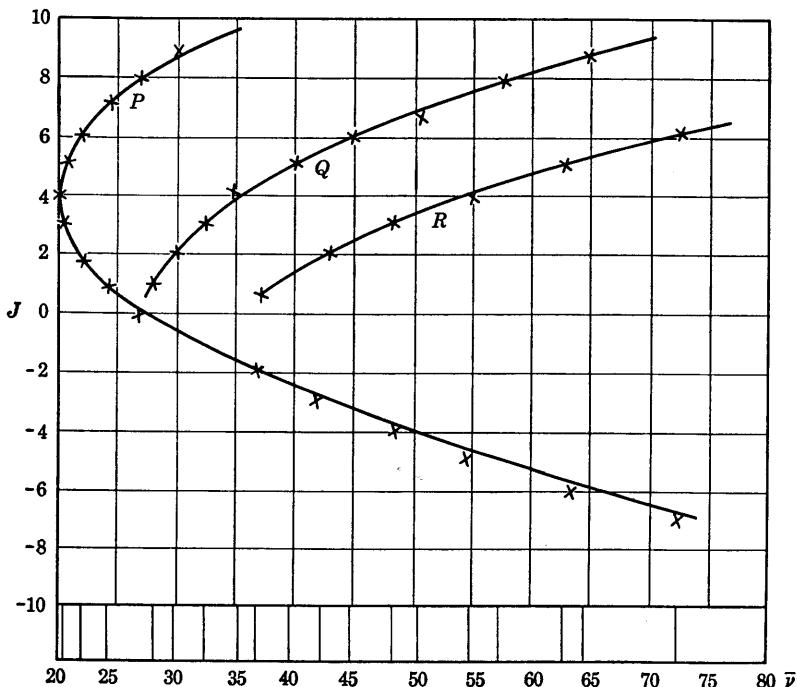


FIG. 5.25. Analysis of the CO band at 5610 Å. (Fortrat diagram.)

minimum wave-number difference occurs for either $\Delta\nu_P$ or $\Delta\nu_R$, and from the figure this is seen to occur for $\Delta\nu_P$ in this band system. By setting the derivative of $\Delta\nu_P$ with respect to J equal to zero, the minimum frequency occurs for

$$J = \frac{1}{2} \frac{R_f^2 + R_i^2}{R_f^2 - R_i^2}$$

From Fig. 5.25, the minimum corresponds to $J = 4$, and hence

$$R_f^2 = 9R_i^2/7$$

If R_i were greater than R_f , the band would degrade to the left and the *band head* (the extreme value of the parabola) would be on the large-rather than the small-wave-number side of the band origin ($J = 0$ intersection). Again, if there were no electronic angular momenta which

could change in the transition, the Q branch would not appear. Another equation for R_f and R_i can be obtained from the first difference of the equation for $\Delta\bar{\nu}_Q$, i.e.,

$$\Delta\bar{\nu}_{Q(J+1)} - \Delta\bar{\nu}_{Q(J)} = \frac{\hbar}{2\pi c M_r} \left(\frac{1}{R_i^2} - \frac{1}{R_f^2} \right) (J + 1)$$

From Table 5.5, $\Delta[\Delta(\Delta\bar{\nu}_{ave})] = 0.8$ and $\frac{\hbar}{2\pi c M_r} \left(\frac{1}{R_i^2} - \frac{1}{R_f^2} \right) = 0.8 \text{ cm}^{-1}$.

From the above equation and $7R_f^2 = 9R_i^2$, R_f and R_i may both be determined: $R_f = 1.3 \times 10^{-10} \text{ m}$ and $R_i = 1.15 \times 10^{-10} \text{ m}$.

TABLE 5.5. Q -BRANCH COMPONENTS OF A CO BAND

J	$\Delta\bar{\nu}_Q$	$\Delta(\Delta\bar{\nu}_Q)$	$\Delta[\Delta(\Delta\bar{\nu}_Q)]$
1	17828.0	1.7	0.7
2	29.7	2.4	0.9
3	32.1	3.3	0.9
4	35.4	4.2	0.8
5	39.6	5.0	0.8
6	44.6	5.8	0.8
7	50.4	6.6	0.9
8	57.0	7.5	
9	64.5		

Influence of Nuclear Spin. The effect of nuclear spin has so far been neglected, and its direct contribution to the energy of a state is inconsequential since the magnetic energy has been seen to be but a small correction to the electrostatic energy. However, in the particular case that the nuclei are identical, the relative orientations of the spins of the nuclei may evidently have an effect upon the symmetry of the wave function describing the molecule as a whole. This is the same situation as for two electrons which are by definition identical, and the relative orientation of their spins determines whether the spatial-coordinate portion of the complete wave function is symmetric or antisymmetric in order that the complete wave function including spin may be antisymmetric. The symmetric and antisymmetric spatial factors represent different energies in general (the symmetric one leading for instance to the homopolar bond). However, the nuclear spin plays a dominant part in the structure of homonuclear molecules through its bearing on symmetry characteristics rather than because of the magnetic energy associated with its orientation.

The over-all diatomic molecular wave function in the coordinate system in which the center of gravity is at rest may be written as the product

of the following five factors, the first two of which refer to the electron structure and the last three to the nuclei:

$$U = u_{eo} \text{ (electron orbital)} \cdot u_{es} \text{ (electron spin)} \cdot u_{nv} \text{ (nuclear vibration)} \\ \cdot u_{nr} \text{ (nuclear rotation)} \cdot u_{ns} \text{ (nuclear spin)} \quad (5.18)$$

In order that U may be antisymmetric in the electrons, the product $u_{eo}u_{es}$ is antisymmetric as has been seen in Sec. 5.3. u_{nv} is simply a function of the separation R and as such is unaffected by an interchange in the nuclei, i.e., it is symmetric. Thus, in accordance with the general principle stated in Sec. 4.4 that wave functions are antisymmetric for elementary particles with spin $\frac{1}{2}\hbar$, the product $u_{nr}u_{ns}$ must have the appropriate symmetry properties. The angular equation for u_{nr} is identical with the angular equation for the hydrogen electron, and therefore the discussion in Sec. 3.4 shows that u_{nr} is symmetric if the nuclear rotational quantum number is even and antisymmetric if it is odd.

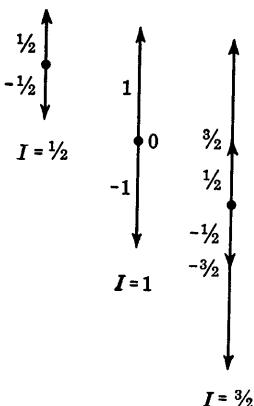


FIG. 5.26. Spin components.

By analogy with the discussion in Sec. 4.6, the function u_{ns} for a homonuclear molecule is the product of the two nuclear spin functions characterizing the possible projections of the spins upon the arbitrarily chosen axis. If I is the total value of the spin of each nucleus in units of \hbar , each may have $(2I + 1)$ possible orientations as

seen by Fig. 5.26 and the number of possibilities for the two together is $(2I + 1)^2$. Our purpose is to see which of these are symmetric and which antisymmetric. To do this, consider the symmetry of linear combinations of the form

$$(m_I)_1(m'_I)_2 \pm (m'_I)_1(m_I)_2$$

where m_I and m'_I are possible projections of the nuclear spins I and the subscripts refer to nuclei 1 and 2. If $m_I = m'_I$ the negative sign leads to zero, and there will be $(2I + 1)$ combinations with a positive sign, i.e., symmetric in spin for an interchange of 1 and 2. If $m_I \neq m'_I$ the product $(m_I)_1(m'_I)_2$ without regard to order can be chosen in $(2I + 1)2I/2$ ways, and the positive sign will be symmetric and the negative antisymmetric for an interchange in 1 and 2; thus there will be $I(2I + 1)$ combinations of each type for $m_I \neq m'_I$. These are seen to exhaust the possibilities since $I(2I + 1) + I(2I + 1) + 2I + 1 = (2I + 1)^2$; $(I + 1)(2I + 1)$ are symmetric and $I(2I + 1)$ are antisymmetric.

From the above discussion it is clear that u_{ns} can be either symmetric

or antisymmetric and the rotational states for which the nuclear rotational quantum number is even and those for which it is odd are separately associated with one or the other depending on whether the nucleus has an even- or odd-half-integer value of \hbar for its spin. This is the same in accordance with present experimental evidence as saying that there are an even or odd number of nucleons (protons and neutrons) in the nucleus. For both protons and neutrons have spins of $\frac{1}{2}\hbar$, and these appear to be the only permanent nuclear constituents that contribute to the spin. As was seen in Sec. 4.4 the wave function of an elementary particle always remains either even or odd, but this rule evidently is not a strict one for a complex of particles such as an atom or a molecule. The spatial symmetry can change if a compensating change takes place in the spin symmetry leaving always a net antisymmetric function for each elementary particle composing the system. Changes of this type, however, are less apt to take place than those in which the spatial and spin symmetries each separately remain the same. In the helium spectrum, for example, the main intercombinations are within the singlet and within the triplet systems, though intersystem combinations do occur involving a reorientation of the relative spin directions. In a similar way a homonuclear molecule may change from u_{ns} symmetric to u_{ns} antisymmetric, but this is a rare event under ordinary circumstances. It is much rarer than intersystem transitions in an electron structure because the magnetic forces upon the nuclear spins are less by a factor of about 1000 than those on electron spins because of the magnitudes of the moments, and it is these forces that would bring about the change u_{ns} (sym.) $\leftrightarrow u_{ns}$ (antisym.).

To summarize the symmetry properties of U , it may be represented as the product of the four critical u 's (omitting u_{nv} which is always symmetric):

$$U = u_{eo}(L_e)u_{es}(S_e)u_{nr}(L_n)u_{ns}(S_n)$$

where L_e , S_e , L_n , and S_n represent the electronic and nuclear orbital and spin quantum numbers, respectively. The electronic function is written as two separate terms (the assumption of Russell-Saunders coupling) though this is not essential, and J , the total angular momentum, is the sum of L 's and S 's. S_n is practically unalterable, certainly in times of the order of optical transitions, and S_e is not so subject to change as the L 's or the relative orientations of these vectors. Thus in changes in J of spectroscopic interest the changes of the L 's are more probable than changes in S 's.

The experimental observations relating to the above are of great interest. Because homonuclear molecules are limited to either even or odd numbers of L_n , half of the rotational levels for each nuclear symmetry

type are missing. In consequence of the selection rule $\Delta J = \pm 1$ (if $\Lambda = 0$), there will then be no rotation or vibration spectra in the red and infrared. There will of course be Raman spectra $\Delta J = 0, \pm 2$ since these are possible in the absence of alternate levels. In the case of electronic transitions there may be rotational structure because, though alternate rotational levels are missing for any one nuclear spin symmetry, these may be the even levels in one electronic state and the odd levels in another electronic state because of opposite electronic orbital angular-momentum parity. Since in general there are different numbers of the molecules with symmetric and antisymmetric nuclear spins, there are different numbers with populated odd and even rotational nuclear states. This leads to alternating intensities in the rotational band structure for homonuclear molecules. H^1H^2 or $O^{16}O^{18}$ are not homonuclear molecules, and the above discussion refers only to such molecules as H^1H^1 or $O^{16}O^{16}$. Mixed molecules such as H^1H^2 do not show alternating intensities.

It is possible to change homonuclear molecules from one nuclear-spin symmetry to another by the introduction of magnetic forces. It is most commonly done by paramagnetic ions having strong local inhomogeneous magnetic fields which provide coupling with the spin angular momenta. In statistical equilibrium one would expect the ratio of the molecules of the two types to be determined by the ratio of the numbers of relative orientations that are symmetric and antisymmetric respectively, i.e., from a previous paragraph, $(I + 1)/I$. Thus for H_2^1 for which $I = \frac{1}{2}$ the equilibrium ratio would be 3:1 for the ortho (symmetric) to the para (antisymmetric) types.

This is the equilibrium concentration generally found to exist, and it is maintained even when the temperature is lowered toward liquefaction. In spite of the abstraction of energy that brings about this lowered temperature, the orthohydrogen remains in its lowest state $I_n = 1$ and does not go over the parahydrogen which can exist in the still lower rotational state $I_n = 0$ unless some catalyst is present such as a paramagnetic salt. If such is present, however, practically all the molecules go over to the para form. If the catalyst is then removed, the temperature may be raised without inducing any change from para to ortho forms and thus practically pure parahydrogen can be prepared and maintained over long periods of time.

In the cases of such molecules as H_2^2 or N_2^{14} the value of I is 1 and the symmetric-antisymmetric ratio is 2:1, again a quantity experimentally observed from the rotational-level intensities. For O_2^{16} for which $I = 0$ there is only one type, all such molecules are symmetric in the nuclear spin (this is obvious since it is zero), and alternate rotational levels are completely absent.

PROBLEMS

5.1. Show that the polarizability of a hydrogenic atom in the ground state may be written

$$\alpha_0 = 16\pi\epsilon_0 a_0^3 \sum_k \left(\frac{k^2}{k^2 - 1} \right)^2 f_k \quad \text{farad-m}^2 \quad \text{where } f_k = \frac{1}{a_0^2} \left(\frac{k^2 - 1}{k^2} \right) \left| X'_{0k} \right|^2$$

Using the functions $Y_{l,m}$ and $R_{n,l}$ given in Sec. 3.4, show that the contribution to α_0 for the states as far as $n = 3$ is

$$\frac{6.59 \times 10^{-41}}{Z^2} \left(\frac{4}{3} \frac{2^{15}}{3^{10}} + \frac{9}{8} \frac{3^6}{2^{13}} \right)$$

Since $\Sigma f = 1$, show that α_0 for atomic hydrogen is given approximately by 9.15×10^{-41} farad-m².

5.2. Using the hydrogenic wave functions of Sec. 3.4, show that the structure factor for the ground state of atomic hydrogen is given by $[1 + (a_{HS}/2)^2]^{-2}$, where $s = (4\pi/\lambda) \sin(\phi/2)$.

5.3. Show that the coherent scattering of X rays by a plane equilateral triangular molecule of BF_3 as a function of angle ϕ with the incident beam is proportional to

$$152 + 60\sqrt{3} \frac{\sin(sl/\sqrt{3})}{sl} + 300 \frac{\sin sl}{sl}$$

where l is the separation of the fluorine atoms, s is $4\pi\lambda^{-1} \sin(\phi/2)$, and λ is the wavelength of the X radiation used.

5.4. The coherent scattering of the K_α X radiation of copper ($\lambda = 1.54 \text{ \AA}$) by carbon tetrachloride shows maxima of intensity at the angles 36° , 67.5° , and 110° with the incident beam. Show that from this data the separation between the chlorine atoms is 3.1 \AA .

5.5. Following the discussion of Sec. 5.1, show that coherent X-ray diffraction maxima would be expected in scattering from a liquid if the atoms of the liquid are on the average separated from their nearest neighbors by a distance l . Given that the smallest value of x satisfying the transcendental equation $\tan x = x$ is $x = 7.72$, show that the relation between the angle ϕ at which the first diffraction maximum is observed in liquid scattering, the wavelength λ , and the average separation l of nearest atoms in the liquid is $\lambda = 1.628 l \sin(\phi/2)$ (Keesom's equation). For the scattering of copper K_α ($\lambda = 1.54 \text{ \AA}$) by mercury, the first diffraction maximum occurs at $\phi = 31^\circ 20'$. Show that the average nearest-neighbor separation is 3.5 \AA .

5.6. The reaction $\text{CH}_4 + 2\text{O}_2 \rightarrow \text{CO}_2 + 2\text{H}_2\text{O}$ is found to liberate 9.12 ev per molecule of CO_2 formed. Using the values for the total energy of binding of the other molecules in this equation which are given in this chapter, show that the total energy of binding of CO_2 is 15.85 ev .

5.7. Show that, if the energy of reaction for a chemical equation is written for conditions of constant atmospheric pressure instead of constant volume, the numerical value must differ by $2.36 \times 10^{-2} \text{ ev}$ for each unit mole difference in the gaseous components on the two sides of the equation. The perfect-gas law is assumed, and volumes of solid or liquid constituents are neglected.

5.8. Consider a tetrahedral molecule such as that of carbon tetrachloride, and show that the moment of inertia of the four chlorine atoms about any axis passing through the central carbon atom is $\frac{8}{3}d_{CCl}^2 M_{Cl}$, where d_{CCl} is the separation between the carbon

and each chlorine atom and M_{Cl} is the mass of a chlorine atom. Such a highly symmetrical structure is known as a spherical top. Given that the separation between chlorine atoms of CCl_4 is 2 Å, calculate the moment of inertia of the molecule, assuming the chlorines are the pure isotope of atomic weight 35.

5.9. The ammonia molecule NH_3 is in the form of a squat triangular pyramid having an equilateral base with the nitrogen atom at the apex of symmetry. Such a molecule is of the symmetrical top type. The distance between the nitrogen atom and each hydrogen atom is 1.014 Å, and the angle between each N—H bond and the axis of symmetry is $83^\circ 15'$. Show that for such a symmetrical geometry the moment of inertia is the same about any axis through the center of gravity perpendicular to the axis of symmetry. Show that the moments of inertia about the axis of symmetry and about an axis perpendicular to the symmetry axis at the center of gravity are 4.98×10^{-47} and 2.60×10^{-47} kg-m², respectively. Show that if the angle between the N—H bond and the axis of symmetry were $52^\circ 7'$ these moments of inertia would be the same and the molecule would have the characteristics of a spherical top.

5.10. The following two sets of orthogonal wave functions represent the bonding electrons for tetravalent carbon (a) for the tetragonal geometry, and (b) for the trigonal geometry. Show how these may be derived from the general forms in Sec. 5.3.

$$(a) \quad \begin{aligned} u_1 &= \frac{1}{2}\{u(S) + \sqrt{3}[u(p_x) + u(p_y) + u(p_z)]\} \\ u_2 &= \frac{1}{2}\{u(S) - \sqrt{3}[u(p_x) + u(p_y) - u(p_z)]\} \\ u_3 &= \frac{1}{2}\{u(S) + \sqrt{3}[u(p_x) - u(p_y) - u(p_z)]\} \\ u_4 &= \frac{1}{2}\{u(S) - \sqrt{3}[u(p_x) - u(p_y) + u(p_z)]\} \end{aligned}$$

$$(b) \quad \begin{aligned} u_1 &= \frac{1}{\sqrt{3}}[u(S) + \sqrt{2}u(p_z)] \\ u_2 &= \frac{1}{\sqrt{6}}[\sqrt{2}u(S) - u(p_x) + \sqrt{3}u(p_y)] \\ u_3 &= \frac{1}{\sqrt{6}}[\sqrt{2}u(S) - u(p_x) - \sqrt{3}u(p_y)] \\ u_4 &= u(p_z) \end{aligned}$$

5.11. Consider the crude approximation of a hydrogen molecule as a rigid structure of two protons and a rigid structure of two electrons held in relative position with respect to one another by an isotropic elastic force having the same force constant as that observed for the vibrational motion of the nuclei. Show that the polarizability α is then given by $4e^2/\omega_c^2 M_r = 2e^2 a^2/U_0$. Using the values of the constants in the text, calculate α for the hydrogen molecule. The dielectric constant of hydrogen at 0°C and a pressure of 1 atm is found to be 1.000264. Given that the dielectric constant is $1 + N'\alpha/\epsilon_0$, where N' is the number of molecules per cubic meter, compare the observed dielectric constant with that which would be calculated on the relatively crude assumptions of this problem.

5.12. Taking $\rho_0 = (\hbar/\omega_c M_r)^{\frac{1}{2}}$ as the amplitude of the zero-point vibration of a molecule, calculate the fractional variation in separation of the H_2 and Cl_2 molecules in their lowest vibrational state.

5.13. A group of absorption lines in the infrared is observed on passing white light through hydrogen fluoride. The measured wave numbers per centimeter of the lines are given below. Derive (a) the angular vibration frequency ω_c for the HF molecule, (b) the zero-point vibrational energy in electron volts, (c) the moment of inertia, and (d) the internuclear separation (neglect anharmonicity).

4203.25	3921.85
4174.69	3879.10
42.85	34.69
11.50	3790.06
4076.14	44.00
39.90	3696.07
02.02	

5.14. Absorption lines with the following wave numbers per centimeter were found by Nielsen in investigating the spectrum of hydrogen iodide:

2300.61	2204.08
2289.34	2189.49
2278.34	2176.82
2266.95	2162.55
2254.71	2148.04
2242.59	2133.29
2217.34	

These constitute the longest wavelength absorption band found for this molecule in the near infrared. Determine the fundamental vibration frequency for this molecule. Calculate the moment of inertia and nuclear separation. Compare the nuclear separations and the restoring forces per unit change in this separation for HF, HCl, and HI using the values given in the text for HCl and from Prob. 5.13 for HF.

5.15. The fundamental vibrational frequency of hydrogen bromide is found to occur at 2649.67 cm⁻¹ wave numbers. What would be the separation in wave numbers of the fundamental vibrational frequencies for hydrogen bromide made up of HBr⁷⁹ and HBr⁸¹? Where would the fundamental frequency of deuterium bromide lie?

5.16. Monochromatic light is scattered by a diatomic molecular substance in which two isotopes are present of mass M_1 and $M_2 = M_1 + \delta M$, where $\delta M/M_1$ is small. Show that because of rotational Raman scattering the scattered spectrum consists of the incident light accompanied by three series of close companion lines, the wave-number separation between members of the three series being given by $2\bar{\nu}_1$, $(2 - \delta M/M_1)\bar{\nu}_1$, and $2(1 - \delta M/M_1)\bar{\nu}_1$, where $\bar{\nu}_1 = \hbar(\pi c M_1 R_0^2)^{-1}$. Find these separations in angstroms for the case of the mercury line at 4046 Å being scattered from chlorine for which the isotopes have the atomic weights 35 and 37.

5.17. Consider the rotation about their center of gravity of two particles having the reduced mass M_r , in the absence of any centrifugal force, which are separated by the distance R_0 . Assuming elastic binding between the particles with a force per unit increase in separation $-k$ and assuming that the angular momentum $R^2 M_r \omega$ is quantized and its square can take only the values $J(J+1)\hbar^2$, show that the separation under the centrifugal stretching force is $R - R_0 = J(J+1)\hbar^2/M_r R^3 k$ and that the energy of rotation is

$$E = \frac{1}{2} \left[\frac{J(J+1)\hbar^2}{M_r R^2} + k(R - R_0)^2 \right] = \frac{1}{2} \frac{J(J+1)\hbar^2}{M_r R_0^2} + \frac{1}{2k} \left[\frac{J(J+1)\hbar^2}{M_r R_0^3} \right]^2$$

if $R - R_0$ is small. Assuming the energy to be given by Eq. (5.9), show that the correction term in rotational energy due to the increase in separation is

$$\frac{a^2}{4U_0} \left[\frac{J(J+1)\hbar^2}{M_r R_0^3} \right]^2$$

5.18. Show that if the term in the first power of ρ is retained in Eq. (5.12), an additional energy term given by the result of the preceding problem occurs in Eq. (5.14).

5.19. Using the expressions given in the text for the wave numbers of the P , Q , and R branches of an electronic band, show that certain sets of differences yield multiples of the initial and final moments of inertia as shown below:

$$\Delta\bar{\nu}_R(J) - \Delta\bar{\nu}_Q(J) = \Delta\bar{\nu}_Q(J+1) - \Delta\bar{\nu}_P(J+1) = \frac{(J+1)\hbar}{2\pi c M_r R_i^2}$$

$$\Delta\bar{\nu}_R(J) - \Delta\bar{\nu}_Q(J+1) = \Delta\bar{\nu}_Q(J) - \Delta\bar{\nu}_P(J+1) = \frac{(J+1)\hbar}{2\pi c M_r R_f^2}$$

Given the following members of these branches in wave numbers per centimeter for a band of the fourth positive group of CO, verify the equations above and determine the initial and final moments of inertia and nuclear separations.

	$\Delta\bar{\nu}_R$	$\Delta\bar{\nu}_Q$	$\Delta\bar{\nu}_P$
$J = 0$	48,338.37		
1	340.94	48,335.00	
2	342.87	333.95	48,328.58
3	345.16	332.64	324.14
4	346.38	330.84	318.84
5	347.17	328.58	313.30
6	347.17	326.15	307.76
7	347.17	322.60	301.69
8	346.38	319.23	295.13

5.20. Using the expressions given in the text for the wave numbers of the P and R branches of an electronic band, show that the following sets of differences between these yield the moments of inertia of the initial and final states:

$$\Delta\bar{\nu}_R(J-1) - \Delta\bar{\nu}_P(J+1) = \frac{\hbar}{\pi M_r R_f^2 c} (J + \frac{1}{2})$$

$$\Delta\bar{\nu}_R(J) - \Delta\bar{\nu}_P(J) = \frac{\hbar}{\pi M_r R_i^2 c} (J + \frac{1}{2})$$

The wave numbers per centimeter of the central 15 lines of the green (O—O) BeO band are given below. Identify the band origin (missing line) and the J numbers of the two branches. Verify the expressions above, and determine the initial and final moments of inertia and the nuclear separations.

21,174.82	21,199.81
178.88	202.88
182.66	205.74
186.41	208.52
189.97	211.12
193.25	213.48

5.21. The following table lists the band origins (after Herzberg) of the vibrational bands for the so-called fourth positive group of carbon monoxide. Using Eq. (5.16) and including a quadratic term in n , show that

$$\bar{\nu}_{n_i+1} - \bar{\nu}_{n_i} = \frac{\omega_{ci}}{2\pi c} [1 + 2\alpha_i(n_i + 1)]$$

$$\bar{\nu}_{n_f+1} - \bar{\nu}_{n_f} = -\frac{\omega_{cf}}{2\pi c} [1 + 2\alpha_f(n_f + 1)]$$

Verify these relations for this group of CO bands, and determine the best mean values for ω_{ci} , ω_{cf} , α_i , and α_f for the states of CO concerned.

$n_i \backslash n_f$	0	1	2	3	4
0	64,703.0†	62,601.8	60,484.7	58,393.2	56,329.4
1	66,231.3	64,087.6	59,881.6	57,818.3
2	67,674.8	65,533.1	63,416.1	61,325.2
3	69,087.8	66,944.3	64,828.1	60,674.5
4	70,469.5	68,323.4	66,199†	64,116.5	62,055.3

† Lesser accuracy.

GENERAL REFERENCES

- Compton, A. H., and S. K. Allison: "X-Rays in Theory and Experiment," D. Van Nostrand Company, Inc., New York, 1935.
- Frisch, O. R.: "Progress in Nuclear Physics," (Neutron Scattering by Cassetts), Academic Press, Inc., New York, 1950.
- Gaydon, A. G.: "Dissociation Energies and Spectra of Diatomic Molecules," Chapman & Hall, Ltd., London, 1947.
- Herzberg, G.: "Infra-red and Raman Spectra of Polyatomic Molecules," D. Van Nostrand Company, Inc., New York, 1947.
- Herzberg, G.: "Spectra of Diatomic Molecules," D. Van Nostrand Company, Inc., New York, 1950.
- Jevons, W.: "Band Spectra of Diatomic Molecules," Physical Society, London, 1932.
- Mott, N. F., and H. S. W. Massey: "Atomic Collisions," Oxford University Press, London, 1933.
- Mott, N. F., and I. N. Sneddon: "Wave Mechanics," Oxford University Press, London, 1948.
- Pauling, L.: "The Nature of the Chemical Bond," Cornell University Press, Ithaca, N.Y., 1940.
- Pauling, L., and E. B. Wilson: "Introduction to Quantum Mechanics," chaps. 10, 12, McGraw-Hill Book Company, Inc., New York, 1935.
- Pearse, R. W. B., and A. G. Gaydon: "Identification of Molecular Spectra," John Wiley & Sons, Inc., New York, 1950.
- Pirenne, M. H.: "Diffraction of X Rays and Electrons by Free Molecules," Cambridge University Press, London, 1946.
- Randall, J. T.: "Diffraction of X-Rays and Electrons by Solids, Liquids and Gases," John Wiley & Sons, Inc., New York, 1934.
- Rice, O. K.: "Electronic Structure and Chemical Binding," McGraw-Hill Book Company, Inc., New York, 1940.
- Sproull, W. T.: "X-Rays in Practice," McGraw-Hill Book Company, Inc., New York, 1946.

CHAPTER 6

STATISTICS OF ATOMIC PROCESSES

6.1. Statistics of Observation. The quantitative description of physical phenomena is formulated in terms of the idealized concepts which are commonly associated with such fields of mathematics as algebra and geometry. But in contrast to mathematics where complete precision of statement is possible, physical science is characterized by observation and experimental comparison in which ultimate precision in a mathematical sense can never be achieved. Geometrical concepts are basic in the description of those spatial relationships and symmetries which are essential to the skeletal structure of a quantitative theory, but the actual objects of observation do not possess the perfection of their idealized prototypes. As a result of certain observations, statements can definitely be made that particular events did or did not occur in certain temporal or spatial intervals. These, however, represent particular instances and not the universally valid association of numbers with phenomena which is necessary for the formulation of a scientific description. The observations upon which physical science is based are primarily those which involve quantitative comparisons of lengths, times, masses, etc., or of pointer readings on dials or scales where the results can be expressed in numbers. It is the association of physical dimensions with the numbers measured that confers scientific significance upon them, but this very association prevents the achievement of any arbitrary precision in their measurement.

The atomic constitution of matter accounts for a basic limit imposed upon precision of measurement. The matter observed and the instruments used for the observation are fundamentally granular, and the limit that is thus imposed on attainable precision cannot be transcended. The present description of elementary particles in terms of a probability amplitude density ψ is not necessarily unique. The concept of continuity may eventually be extendable into the region of atomic and nuclear dimensions, but it appears much less probable that quantitative physical measurement will be extended much beyond the limits that have now in many cases been reached. Here the very act of making the measurement itself disturbs the system being measured to such an extent that

it is not possible to make a precise statement as to its subsequent behavior. These measurements which impinge upon the ultimate atomic limits are very informative concerning atomic and nuclear properties. However, such measurements require statistical analysis for their evaluation.

The concepts of randomness and probability are of particular usefulness in describing such observations. Randomness may be associated with different causes even though the same statistical methods may be used to analyze the different situations. One type of randomness may be associated with the complexity of initial and boundary conditions of large-scale systems which could in principle be described in sufficient detail to allow precise prediction. This randomness is due either to imprecise measurement of the initial conditions or to the limitations of subsequent calculation, so that the future course of the system is not well known. An example of this type of situation is the flipping of a coin. Precise measurement or control of the initial and boundary conditions would allow accurate prediction of whether the coin would fall heads or tails. Nevertheless, in actual practice, the operation is purely random and the appearance of heads is statistical. Another type of randomness is due to the discreteness of atomic particles and the limitation on measurement described by the uncertainty principle. An example of this is the randomness of the instant of time when identical excited atoms radiate. It is unprofitable to speculate on the possible details, which we cannot at present observe but which might be construed as affecting the initial identity of the atoms. The uncertainty relations indicate the limits on these details beyond which our present concepts do not extend. A single measurement made upon a system which exhibits random behavior due to either of these causes is of little utility. Nevertheless ensembles of these identical systems can be characterized by average properties which can be measured, and these average parameters can be used to predict average behaviors.

Probability, then, can be used to describe the average chance that a particular measurement will indicate a certain result, in the sense that a large number of repetitions of the measurement upon similar systems will indicate the certain result in a fraction of the trials equal to the normalized probability. The ultimate justification of the basic hypotheses and applicability of statistical techniques to atomic phenomena must be the concordance of the resulting description with experimental observation. The theory will be considered successful and adequate, in accordance with the experimental point of view here adopted, to the extent that its quantitative predictions are in agreement with experimental observation with due regard to the fact that such agreement will be of a statistical type and will not represent the ultimate in precision in individual instances. For a discussion of the various troublesome ideo-

logical and logical questions involved in probability theory, reference should be made to treatises on this subject.

As an illustration of the statistical approach to measurement, consider the use of a counting device for determining the rate of occurrence of atomic events such as the scattering of alpha particles. The record of the number of counts may be portrayed as a number on a register and recorded at regular intervals separated by the time τ . The data of observation are then the series of numbers indicating the number of events occurring before the successive time instants. Let n_i be the increment in the number portrayed by the register during the i th interval τ . If the series of numbers n_i corresponds to a certain type of pattern, called that of normal error, the average of the numbers is considered to represent a significant physical measurement. Previous experience with measurements of all kinds has indicated that such a pattern is characteristic of a situation free of secular trends or controllable fluctuating factors affecting the results but irrelevant to the particular phenomenon being measured. This pattern is the hypothetical chance distribution of numbers based upon a postulated a priori probability as an arbitrary parameter. Thus if p is the assumed probability of observing a count in the small interval δt determined by the resolving time of the apparatus (no more than one count can possibly occur in δt), the chance of observing exactly n counts in m intervals ($m = \tau/\delta t$) is the term in the expansion of $[p + (1 - p)]^m$ for which p occurs to the n th power and $(1 - p)$ to the $(m - n)$ th power, i.e.,

$$P_n = \frac{m!}{(m - n)!n!} p^n(1 - p)^{m-n} \quad (6.1)$$

The factor p^n is the probability that each n specified intervals will contain one count, the factor $(1 - p)^{m-n}$ is the probability that each of the remaining $(m - n)$ specified intervals will not contain a count, and the factorial factor is the number of ways of dividing up the m intervals into n of one type (one count) and $(m - n)$ of the other (zero counts). Equation (6.1) describes the error pattern referred to above.

The degree of concordance between any set of observations, such as that of the numbers of counts found in successive periods of length τ , and Eq. (6.1) may be determined graphically or, often more conveniently, by comparing the deviations arithmetically. For this purpose it is useful to start from the identity

$$[px + (1 - p)]^m = \sum_{n=0}^m \frac{m!}{(m - n)!n!} p^n(1 - p)^{m-n}x^n = \sum_{n=0}^m P_n x^n$$

If x is set equal to 1, the equation becomes $\Sigma P_n = 1$ which states that the sum of the probabilities of all possible alternatives is unity. Differ-

differentiating the identity once with respect to x ,

$$mp[px + (1 - p)]^{m-1} = \sum_{n=0}^m nP_n x^{n-1}$$

Letting $x = 1$, the equation $mp = \Sigma nP_n$ is obtained and the sum is by definition the *average value* of n which will be written \bar{n} , i.e.,

$$\bar{n} = \Sigma nP_n = mp$$

Differentiating the identity twice with respect to x and setting $x = 1$ in the result,

$$m(m - 1)p^2 = \sum_{n=0}^m n(n - 1)P_n$$

and, using the definition of \bar{n} ,

$$\sum n^2 P_n = \bar{n}^2 + \left(\bar{n} - \frac{\bar{n}^2}{m} \right)$$

The above sum is the average value of n^2 , written $\overline{n^2}$. The *standard deviation*, generally written σ , is defined as

$$\sigma = (\overline{n^2} - \bar{n}^2)^{\frac{1}{2}} = \left[\bar{n} \left(1 - \frac{\bar{n}}{m} \right) \right]^{\frac{1}{2}} \quad (6.2)$$

Higher-order deviations can be calculated in the same way and algebraic comparisons made with the data of observation.

In many applications of statistical methods, all the numbers m , n , and $m - n$ are very large. In such cases Stirling's approximation

$$\ln n! \cong n \ln n - n$$

is adequate for acceptable precision in P_n and the factorial expression of Eq. (6.1) becomes an analytic function. In this case the general form of P_n in the neighborhood of $n = \bar{n} = mp$ can be found by setting

$$n = \bar{n}(1 + s)$$

where s is assumed to be small compared to unity. Expanding P in terms of s ,

$$\begin{aligned} \ln P_s &= m \ln m - \underbrace{(m - \bar{n} - \bar{n}s) \ln (m - \bar{n} - \bar{n}s)}_{-\bar{n}s^2} \\ &\quad - (\bar{n} + \bar{n}s) \ln (\bar{n} + \bar{n}s) + \bar{n}(1 + s) \ln \frac{\bar{n}}{m} \\ &\quad + (m - \bar{n} - \bar{n}s) \ln \frac{m - \bar{n}}{m} \\ &\cong -\frac{\bar{n}s^2}{2(1 - \bar{n}/m)} = -\frac{(\bar{n}s)^2}{2\sigma^2} \end{aligned}$$

where $\ln(1+s) \cong s - s^2/2$ and Eq. (6.2) have been used. If δt is sufficiently short that $p = \bar{n}/m$ is very small in comparison with unity, σ^2 is approximately equal to \bar{n} and P_s may be written

$$P_s = P_0 e^{-(\bar{n}s)^2/2\sigma^2} \cong P_0 e^{-(n-\bar{n})^2/2\bar{n}} \quad (6.3)$$

Equation (6.3) is the so-called *normal-error curve* or Laplace distribution law and is plotted in Fig. 6.1. If the observed values of n_i are clustered about the mean value of n_i in the symmetrical way in which the ordinates of this curve are clustered about the ordinate $n = \bar{n}$, the distribution is said to be *normal*. If the distribution of experimental values about their mean corresponds to Eq. (6.1), or to Eq. (6.3) in the approximation

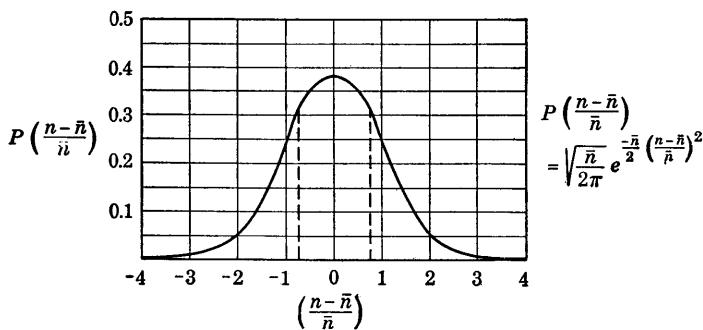


FIG. 6.1. Normal-error curve.

where it applies, the average value of the n_i 's is considered to be a reliable measure of the number of counts in the interval τ . The fractional probable error, which is a measure of the precision of a single observation, is that value of $(n - \bar{n})/\bar{n}$, say $(n_p - \bar{n})/\bar{n}$, for which the area under the curve of P as a function of $(n - \bar{n})/\bar{n}$ between $\pm (n_p - \bar{n})/\bar{n}$ is equal to the area lying on either side beyond these points. The curve falls to a fraction f of its peak value when $(n - \bar{n})/\bar{n} = [2 \ln(1/f)/\bar{n}]^{1/2}$; thus the fractional probable error is proportional to $\bar{n}^{-1/2}$. Integration of P from $-\infty$ to $+\infty$ yields the numerical constant of proportionality as 0.675. The standard deviation σ , which is approximately equal to $\sqrt{\bar{n}}$, is a very convenient definition of statistical precision, for if, say, 10,000 counts are observed in a given time interval the standard deviation, which is a measure of the precision with which 10,000 may be taken as the correct measurement, is seen to be $10,000^{-1/2}$, or 1 per cent of 10,000, and within the limitations of the assumptions involved the reliability of the measurement can be immediately assessed. For a value of \bar{n} which is not large compared to 1, the Stirling approximation does not hold and the expected distribution curve is skewed because of the limitation of possible values on the zero side of \bar{n} . The theoretical curve to be used in place of Eq. (6.3)

for small n is

$$P(n) = \frac{(\bar{n})^n e^{-\bar{n}}}{n!} \quad (\text{see Prob. 6.1 and Fig. 6.2})$$

In the illustration it should be noted that what has been measured is a number of counts per unit time, which is not yet of physical significance. These counts may be due to several causes. To determine the number due to alpha particles scattered from the foil, measurements must be made with the foil in place and with it removed, and the background counts subtracted from the total. Also, the possible effects of geometry, neighboring objects, etc., must be tested by suitably varying the experimental arrangements.

The preceding discussion is illustrative of the statistical method of approach to a particular problem of measurement. It has validity only to the extent that the observed distribution of the observations n about the average \bar{n} resembles closely the normal-error law [Eq. (6.2)] and that the inference from this that the only errors involved are those of chance is justifiable. The various terms used are not precisely defined but merely statistically defined. In consequence, the conclusions derived from the theory are subject to a statistical uncertainty, and the approach is without the satisfactory ideological foundation that characterizes strict mathematical statements. However, experience has shown that the statistical approach provides results that are reliable within the limitations that are clearly implied, and, as random fluctuation is a basic factor in all measurement, statistical limitations must be accepted to a greater or lesser degree in all experimental science.

6.2. Thermal Equilibrium. The central problem of macroscopic physics is the description of the large aggregates of atoms which constitute matter in bulk in terms of the properties which characterize the individual atomic components. The general problem of formulating an adequate and comprehensible theory of all large-scale phenomena in these terms is of course very formidable indeed and far from any complete solution. The statistical approach, however, is a very promising one and not only provides considerable insight into the properties of aggregates far too numerous to study in any other way but also yields many results of great

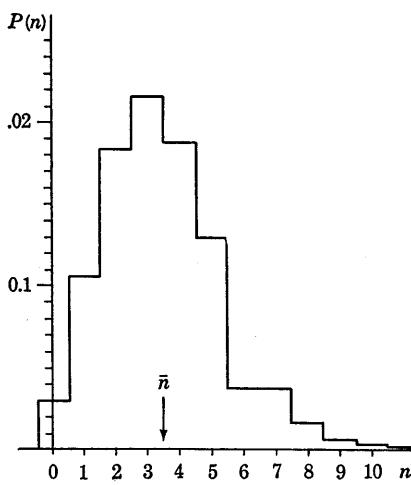


FIG. 6.2. A Poisson distribution curve.

practical importance concerning the behavior of matter in bulk. Only the simplest types of atomic aggregation can as yet be handled even in this way, and the present discussion is limited to those stationary or quasi-stationary situations which display that degree of constancy and uniformity of observable properties known as *thermal equilibrium*. The general statistical approach to the description of systems in thermal equilibrium with one another does not depend critically upon the characteristics of the individual component systems. The form of the statistics and the significant parameters in terms of which the description is expressed depend of course on the properties of the systems concerned. But the nature of the method is dependent simply upon the ability of the constituent systems to be described in terms of the energy associated with each of them and the existence of a mechanism whereby energy can be transferred from one to another. For purposes of illustration the aggregate can be considered as an assembly of familiar large-scale mechanical or electric systems such as springs, pendulums, confined gases, electric circuits, etc., loosely coupled together so that both the energy of each and the sum of these, constituting the energy of the total assembly, may be sufficiently well defined for the purpose of the description of the physical characteristics that the assembly exhibits.

A logical approach to the statistical description of aggregates is that of Gibbs, who introduced the concept of an *ensemble* which is a hypothetical assembly of systems with certain idealized properties with which actual physical assemblies of systems can be identified to a greater or lesser degree depending on the extent to which the idealized properties correspond to those that are physically observed. If this correspondence appears to be close, confidence can reasonably be placed in the applicability to the assembly of physical systems of the conclusions deduced from the idealized hypothetical ensemble. In this sense the principle is analogous to that of comparing actual observed deviations from the mean with the normal-error curve and, if the correspondence is found acceptable, assuming that the errors are those of chance and that the reliability of the set of observations is measured by $\bar{n}^{-\frac{1}{2}}$. An ensemble is an indefinitely large number of identical physical systems that are otherwise completely isolated from the rest of the universe. The systems may be any physical entities capable of storing and exchanging energy such as weights and springs, oscillatory electric circuits, gases in elastic containers, etc. These are possessed of a given amount of energy which can be interchanged between the constituent systems by means of interactions which are, however, not so strong as to preclude the association of a certain fairly definite energy content with each system separately.

An individual system is characterized by the amount of energy, or fraction of the total energy U of the ensemble, that it possesses. This

of course cannot be determined with arbitrary precision. It is assumed that Δu is the smallest energy difference that can be measured. In terms of Δu as a unit, an energy scale $u_1, u_2, \dots, u_j, \dots$ can be set up, where $u_{j+1} - u_j = \Delta u$. The ensemble can then be completely described for statistical purposes by stating the number of systems which have energies within the intervals on the energy scale, i.e.,

$$\begin{aligned} n_1 &\text{ systems have the energy } u_1 \\ n_2 &\text{ systems have the energy } u_2 \\ \cdot & \quad \cdot \\ n_j &\text{ systems have the energy } u_j \\ \cdot & \quad \cdot \end{aligned}$$

There are of course the conditions

$$\Sigma n_j = N, \quad \Sigma u_j n_j = U \quad (6.4)$$

where N is the total number of systems and U is the total amount of energy. It is further postulated for the moment that all the systems of the ensemble are of the same type and that there is no weighting factor differentially favoring the occurrence of a system in one region of the energy scale over another. Occupancy by a system of any interval Δu throughout the available energy scale is equally likely, and the probability of finding any particular system in a range Δu is some small quantity δp equal to the ratio of Δu to the available energy range. Many different particular distributions of the individual systems lead to the same state of the ensemble as specified by the numbers of systems with particular energy values. If all particular distributions are equally likely by hypothesis, the probability that the ensemble shall be characterized by n_1 systems with energy u_1 , n_2 with energy u_2 , etc., is equal to the number of particular distributions so described or the number of ways of choosing subgroups composed of n_1, n_2, n_3 , etc., members from the total group N of systems in the ensemble times δp raised to the N th power. The factor multiplying $(\delta p)^N$ is seen to be the number of ways of choosing N systems divided by the number of permutations within the subgroups n_1, n_2, n_3 , etc., since these permutations are without influence on the description of the ensemble. Thus the likelihood of the description n_1 systems of energy u_1, n_2 of energy u_2 , etc., is proportional to

$$\frac{N!}{n_1! n_2! \cdots} (\delta p)^N = \frac{N!}{\prod_j n_j!} (\delta p)^N = P(\delta p)^N \quad (6.5)$$

where P is defined as $N! \left(\prod_j n_j! \right)^{-1}$.

The problem of finding the most likely description of the ensemble is then that of maximizing Eq. (6.5) in a way consistent with the conditions of Eqs. (6.4) [since here $(\delta p)^N$ is a constant this condition simplifies to maximizing P]. This can be done most conveniently by first taking the logarithm of Eq. (6.5). Since by hypothesis all the quantities N and n_j are indefinitely large, Stirling's approximation applies without question.

$$\ln P = N \ln N - N - \Sigma n_j \ln n_j + \Sigma n_j \quad (6.6)$$

Then, applying the method of Lagrangian undetermined multipliers, Eqs. (6.4) are multiplied by arbitrary constants λ and μ , respectively, and added to Eq. (6.6).

$$\ln P + \lambda N + \mu U = N \ln N - N - \Sigma n_j \ln n_j + (\lambda + 1) \Sigma n_j + \mu \Sigma u_j n_j$$

The second and third terms on the left and the first and second terms on the right are constants. Now assume a small change δn_j is made in each n_j which induces a change $\delta(\ln P)$ in $\ln P$.

$$\begin{aligned} \delta(\ln P) &= -\Sigma \ln n_j \delta n_j - \Sigma \delta n_j + (\lambda + 1) \Sigma \delta n_j + \mu \Sigma u_j \delta n_j \\ &= -\Sigma (\ln n_j - \lambda - \mu u_j) \delta n_j \end{aligned}$$

The δn_j 's are not all independent because of the two conditions [Eq. (6.4)]. However, all but say δn_1 and δn_2 are independent. Thus an extreme value of $\ln P$ exists if $\delta(\ln P) = 0$, or if

$$\Sigma (\ln n_j - \lambda - \mu u_j) \delta n_j = 0$$

If $\lambda - \ln n_1 + \mu u_1$ and $\lambda - \ln n_2 + \mu u_2$ are arbitrarily set equal to zero so that δn_1 and δn_2 can have any values and these terms disappear from the sum above, then the sum of the remaining terms must vanish for arbitrary choices of the δn_j 's, and the only way for this to be so is for all of the parentheses to vanish separately, i.e.,

$$\ln n_j = \lambda + \mu u_j$$

or

$$n_j = e^\lambda e^{\mu u_j} \quad (6.7)$$

Equation (6.7) is known as the *Boltzmann distribution law*, and it completely specifies the condition of the ensemble of systems to within the two arbitrary constants λ and μ . The condition $\delta(\ln P) = 0$ clearly represents a maximum as can be seen on examining the behavior of the denominator $\prod_j (n_j!)$ which has a smallest value when the n_j 's are all equal. The degree of likelihood of the distribution given by Eq. (6.7) as compared with any other can be ascertained by considering $\ln P$ of Eq. (6.6) as a function of n_j in the neighborhood of the maximum value.

Expanding $\ln P$ in a Taylor series,

$$\ln P = (\ln P)_{\max} + \left[\frac{\partial}{\partial n_j} (\ln P) \right]_{\max} \delta n_j + \frac{1}{2} \left[\frac{\partial^2}{\partial n_j^2} (\ln P) \right]_{\max} \delta n_j^2 + \dots$$

By virtue of the fact that $\ln P$ is maximized, $[\partial(\ln P)/\partial n_j]_{\max}$ vanishes and

$$\frac{\partial^2}{\partial n_j^2} (\ln P) = - \frac{\partial}{\partial n_j} (\ln n_j) = - \frac{1}{n_j}$$

so

$$P = P_m e^{-\frac{n_j}{2} \left(\frac{\delta n_j}{n_j} \right)^2} \quad (6.8)$$

to this approximation. As the n_j 's of an ensemble are indefinitely great, the probability of any distribution for which the fractional deviation

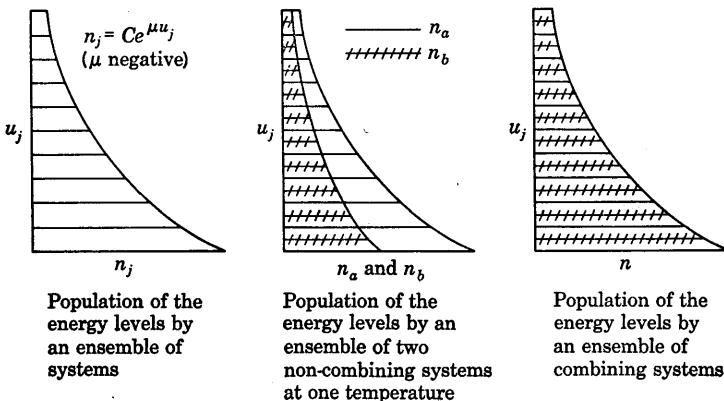


FIG. 6.3. Representative energy-level populations.

$\delta n_j/n_j$ is not vanishingly small is negligible. Thus Eq. (6.7) (illustrated in Fig. 6.3) may be taken as representing without any uncertainty the distribution of systems in energy. If there is a finite number of systems in an actual physical assembly, there would of course be a finite probability of finding a departure from the Boltzmann distribution law.

The limitation on the preceding argument that the systems in the ensemble are all exactly the same is not essential. The physical description of the energy states may, for instance, fall into several separate categories, such as translation, rotation, vibration, electronic excitation, etc. Thus if α_k represents the k th rotational state of a system and β_l the l th state of vibratory motion, the total energy of these two types is $u_j = \alpha_k + \beta_l$. It is convenient to write the sum $\Sigma e^{\mu u_j}$, which occurs very frequently, as the letter Z and give it the name *partition function*. Thus

$$Z = \Sigma e^{\mu u_j}$$

and in terms of the partition function the sum of Eq. (6.7) over j is

$$N = \Sigma e^\lambda e^{\mu u_i} = e^\lambda Z$$

and

$$n_j = \frac{N}{Z} e^{\mu u_i}$$

In terms of the categories α and β

$$n_j = n_{k,l} = N \frac{e^{\mu(\alpha_k + \beta_l)}}{\sum_{k,l} e^{\mu(\alpha_k + \beta_l)}}$$

If this quantity is now summed over either k or l ,

$$n_k = \sum_l n_{kl} = N \frac{e^{\mu\alpha_k} \sum_l e^{\mu\beta_l}}{\sum_k e^{\mu\alpha_k} \sum_l e^{\mu\beta_l}} = N \frac{e^{\mu\alpha_k}}{\sum_k e^{\mu\alpha_k}} = \frac{N e^{\mu\alpha_k}}{Z_k}$$

and analogously for n_l . Thus the distribution of systems over the energy states of the form α_k is the same as the distribution over those of the form β_l quite independently of the other form of energy.

The ensemble may also be composed of two quite distinct types of systems, as for example two types of molecules a and b which do not react chemically with one another. If this were the case, the P analogous to that of Eq. (6.5) would be

$$P = N_a! \left(\prod_j n_{aj}! \right)^{-1} N_b! \left(\prod_k n_{bk}! \right)^{-1}$$

The equations of constant numbers of each type and a common constant energy content would be

$$\sum_j n_{aj} = N_a, \quad \sum_k n_{bk} = N_b, \quad \sum_j n_{aj} u_{aj} + \sum_k n_{bk} u_{bk} = U$$

Proceeding exactly as before, multiplying the first of these by λ_a , the second by λ_b , and the third by μ , and setting $\delta(\ln P)$ equal to zero subject to the equations of condition, it is seen that

$$n_{aj} = e^{\lambda_a} e^{\mu u_{aj}} = \frac{N_a}{Z_a} e^{\mu u_{aj}} \quad (6.7')$$

and

$$n_{bk} = e^{\lambda_b} e^{\mu u_{bk}} = \frac{N_b}{Z_b} e^{\mu u_{bk}}$$

Again the distribution laws are of the same form as Eq. (6.7), and the constant μ is common to them both. If the systems in the ensemble were fundamentally of the same type but capable of taking on two different physically recognizable forms (as in the case of atoms which could exist singly or in combinations of diatomic molecules), the argument is again very similar. The P analogous to Eq. (6.5) would be

$$P = \frac{N!}{N_a!N_b!} \frac{N_a!}{\prod_j n_{aj}!} \frac{N_b!}{\prod_k n_{bk}!}$$

and the equations of condition would be

$$\sum_j n_{aj} + \sum_k n_{bk} = N, \quad \sum_j n_{aj}u_{aj} + \sum_k n_{bk}u_{bk} = U$$

Proceeding as before, using the constants λ and μ and setting $\delta(\ln P)$ equal to zero subject to the equations of condition, it is seen that

$$n_{aj} = e^{\lambda} e^{\mu u_{aj}}, \quad n_{bk} = e^{\lambda} e^{\mu u_{bk}} \quad (6.7'')$$

Again the distribution laws for the two forms of the systems are the same as Eq. (6.7). From the first equation of condition,

$$e^{\lambda} \left(\sum_j e^{\mu u_{aj}} + \sum_k e^{\mu u_{bk}} \right) = e^{\lambda} (Z_a + Z_b) = N$$

Thus e^{λ} may be written in terms of the partition functions and N . Also

$$\begin{aligned} N_a &= \sum_j n_{aj} = e^{\lambda} Z_a = N \frac{Z_a}{Z_a + Z_b} & \frac{N_a}{N_b} &= \frac{Z_a}{Z_b} \\ N_b &= \sum_k n_{bk} = e^{\lambda} Z_b = N \frac{Z_b}{Z_a + Z_b} \end{aligned}$$

Thus the ratio of the numbers of systems in the two forms is equal to the ratio of the partition functions corresponding to the forms. The above discussion shows that the complexity of the systems forming the ensemble is without influence on the form of the distribution law. This is the same for the different physical forms that the energy may take and for different types of systems sharing this energy provided only that the energy is shared and the constant μ is the same in all cases. μ will be related to the temperature, and the constancy of μ is equivalent to uniformity of temperature as the physical criterion for thermal equilibrium.

The partition function Z defined a few paragraphs above is a particularly convenient quantity in terms of which to establish the connec-

tion between the statistics of the ensemble and the laws of thermodynamics. Thus

$$n_j = N \frac{e^{\mu u_j}}{\sum_j e^{\mu u_j}} = \frac{N}{\mu} \frac{\partial}{\partial u_j} (\ln Z) \quad (6.9)$$

and

$$\bar{u} = \frac{U}{N} = \frac{\sum_j u_j e^{\mu u_j}}{\sum_j e^{\mu u_j}} = \frac{\partial}{\partial \mu} (\ln Z)$$

These expressions suggest the connection between the ensemble here under discussion and thermodynamics. For since Z is a function of μ and the u_j 's,

$$\begin{aligned} d(\ln Z) &= \frac{\partial}{\partial \mu} (\ln Z) d\mu + \sum_j \frac{\partial}{\partial u_j} (\ln Z) du_j \\ &= \bar{u} d\mu + \frac{\mu}{N} \sum_j n_j du_j \end{aligned}$$

If the product of $\mu \bar{u}$ is subtracted from $\ln Z$,

$$d(\ln Z - \mu \bar{u}) = -\mu \left(d\bar{u} - \frac{1}{N} \sum_j n_j du_j \right)$$

$d\bar{u}$ is the change in average internal energy of a system [δU of Eq. (1.7)] and the sum $\frac{1}{N} \sum_j n_j du_j$ represents the average work done on a system

in altering the external parameters upon which du_j 's depend; the amount of work done upon the n_j systems if the energy of each is changed by du_j being $n_j du_j$. The du_j 's would correspond for instance in the case of macroscopic systems to tightening springs, separating condenser plates, or compressing elastic volumes. Thus $-\frac{1}{N} \sum_j n_j du_j$ equals the average work done by the system [δW of Eq. (1.7)]. Then $-(1/\mu) d(\ln Z - \mu \bar{u})$ can be identified with δQ or $T dS$ of Eq. (1.8).

$$\begin{aligned} T ds &= \delta Q = \delta U + \delta W \\ -\frac{1}{\mu} d(\ln Z - \bar{u}\mu) &= d\bar{u} - \frac{1}{N} \sum_j n_j du_j \end{aligned}$$

T of Eq. (1.8) is in degrees absolute whereas μ is of the dimensions of reciprocal energy, so a dimensional constant of proportionality is required

in the identification of μ with T , that is, $\mu = -1/kT$, where k is this constant. The entropy per system may then be identified with k times $(\ln Z - \mu\bar{u})$ or

$$S = k \ln Z + \frac{\bar{u}}{T} \quad (6.10)$$

The other thermodynamic functions can immediately be written in terms of $\ln Z$; thus

$$\psi = \bar{u} - TS = -kT \ln Z, \quad \Phi = \bar{u} - TS + pV = pV - kT \ln Z$$

Using Eqs. (6.6), (6.7), and (6.9), it can readily be shown that the partition function Z is related to the probability P by

$$\ln P = N \left(\ln Z + \frac{\bar{u}}{kT} \right) \quad (6.11)$$

so that the entropy per system S can equally well be written

$$S = \frac{k}{N} \ln P \quad (6.10')$$

The conclusion of the preceding discussion is that the statistical method of approach to the problem of the distribution of energy over an ensemble of systems provides a consistent description which is concordant with the statements of thermodynamics which are in turn of established applicability to all large-scale physical phenomena. The statistical results, however, go considerably beyond the statements of thermodynamics in that a detailed description is furnished of the way in which the energy is divided among the systems making up the assembly in thermal equilibrium:

$$n_j = e^\lambda e^{-u_j/kT} \quad (6.12)$$

The constants λ and k are not given by the general theory but must be determined by comparison with observation. It should be mentioned that Eq. (6.12) is an exact description only for an ensemble as previously defined. If the actual physical assembly contains only a finite number of systems, deviations from the distribution law would be expected in accordance with Eq. (6.8). Also, if the systems of the ensemble do not have the properties assumed for them, if for instance they are not distinguishable macroscopic systems, Eq. (6.12) cannot be assumed to apply. These points will be considered in subsequent sections.

6.3. Applications of Classical Statistics. The subject of thermodynamics, which lies outside the scope of atomic physics, is itself the broadest area of application of the statistical treatment of mechanical systems.

Equations (6.10) and (6.12) are the basis of the statistical foundation of thermodynamics, and the entire field of physical chemistry is the application of thermodynamics to physical and chemical processes. However, there are certain other areas of fundamental processes which are concerned with the partition of energy among systems in a physical assembly and the variations in this partition which are not generally included as part of chemical thermodynamics. These applications are of great importance in themselves and also excellent illustrations of the application of the statistics to physical phenomena.

The energy u_j is generally described by some continuous parameter such as the velocity of a center of gravity, the extension of a mechanical oscillator, the potential difference between the plates of a condenser, or the change in volume of an elastic gas container. The mean energy of a system in thermal equilibrium with its surroundings can be expressed in terms of this parameter which will be written as q . A quite general assumption that includes many cases of interest is that $u_j = aq^n$. In this case, assuming that the systems are weighted proportionally to dq , we have

$$\bar{u} = \frac{\int_0^\infty aq^n e^{-aq^n/kT} dq}{\int_0^\infty e^{-aq^n/kT} dq}$$

This is the definition of an average of the quantity aq^n if the probability of observing a particular system with the energy in the range from q to $q + dq$ is $e^{-aq^n/kT} dq$. The numerator can be integrated immediately by parts to give

$$\begin{aligned} \bar{u} &= \frac{\left[-\frac{kT}{n} q e^{-aq^n/kT} \right]_0^\infty + \frac{kT}{n} \int_0^\infty e^{-aq^n/kT} dq}{\int_0^\infty e^{-aq^n/kT} dq} \\ \bar{u} &= \frac{kT}{n} \end{aligned} \quad (6.13)$$

Thus the mean thermal energy associated with any parameter which enters the expression for the energy of the system as aq^n is kT/n .

Consider for instance the case of a suspension of small particles in a fluid as discussed in Sec. 2.2. The energy u_j is the gravitational potential energy, or the work done against gravity in raising the particle from some fiducial level h_1 to a level h_2 . This is $m'g(h_2 - h_1)$, where m' is the difference between the mass of the particle and that of the fluid displaced. The ratio of the number of particles at or near h_2 to the fiducial

number at h_1 is then

$$\frac{n_2}{n_1} = e^{-m'g(h_2-h_1)/kT}$$

This equation is known as the law of atmospheres and is of very general applicability to a statistical system in a uniform force field such as that of gravity near the surface of the earth. If h is measured from the level at which the number of particles per unit volume is n'_0 ($n_1 = n'_0, h_1 = 0$), then $n' = n'_0 e^{-m'gh/kT}$. The change in vertical force per unit area due to the suspended particles in dh , which is the change in pressure due to them, is seen from Fig. 6.4 to be $-dp = m'gn' dh$. Taking the variation of n' with h from the atmosphere law, $dn' = (-m'g/kT)n' dh$ or $dp = kT dn'$. Writing dn' as dn/V , where V is the volume occupied by n particles at any level, the equation may be integrated to yield $pV = nkT$ taking the integration constant as zero. Thus the partial-pressure law for such a suspension is the perfect-gas law under our assumptions.

The energy of a simple harmonic oscillator such as a mass m on a spring of restoring constant s is $\frac{1}{2}m\dot{x}^2 + \frac{1}{2}sx^2$, where x is the coordinate representing the displacement of m from its equilibrium position where it would have zero energy at rest. In accordance with the foregoing theory this system will be in constant small oscillation if it is in equilibrium with its surroundings at a temperature T . With each parameter x and \dot{x} will be associated an amount of energy $\frac{1}{2}kT$, that is,

$$\bar{u}_{\text{kin}} = \frac{1}{2}\bar{m}\dot{x}^2 = \frac{1}{2}kT = \bar{u}_{\text{pot}} = \frac{1}{2}sx^2$$

or $\sqrt{\bar{x}^2} = \sqrt{kT/s}$ and $\sqrt{\dot{x}^2} = \sqrt{kT/m}$. Though the average values of both x and \dot{x} are zero, the root-mean-square values do not vanish and from a measurement of them for any system the value of the universal constant k can be determined.

The variations in displacements, velocities, etc., due to the mean energy of motion of a system in equilibrium with its surroundings at a temperature T impose a fundamental limitation upon the precision of measurement. The only way in which this limitation can be lessened is by lowering the temperature, and this is frequently resorted to in cases where the necessary precision requires it. Complicated multiply periodic or aperiodic systems in which there are dissipative elements require a more extended analysis. Dissipative elements convert ordered

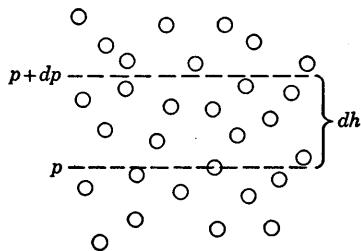


FIG. 6.4. Construction for the derivation of the law of atmospheres.

kinetic and potential energy into heat and, necessarily, heat back into ordered energy, since equilibrium is assumed. Such systems contain damping or resistive elements, and these are coupled with the heat reservoir in such a way that they are essentially transformers dissipating ordered energy and also recovering thermal energy back into mechanical or electrical form.¹⁻³

As an example of a multiply periodic system (following an analysis due to Nyquist⁴), consider a long parallel-wire transmission line shorted at both ends as in Fig. 6.5. A system of standing electromagnetic waves can exist in such a one-dimensional system for which the potential difference is zero at the ends, that is, $\lambda/2 = l/n$ where l is the length of the wires, λ is a wavelength, and n is an integer.

Since the velocity of propagation is c , approximately the velocity of light, the characteristic frequencies of

Fig. 6.5. Standing waves on a parallel-wire line.

are $\nu = (c/2l)n$. If n is large so that ν can be considered an essentially continuous function of n , the number of waves in the frequency range $d\nu$ is $dn = (2l/c) d\nu$. The electric and magnetic energies associated with a wave are quadratic functions of both the electric and magnetic fields, so the thermal energy associated with each characteristic wave is kT and the thermal energy of the system in the frequency range $d\nu$ is $(2kTl/c) d\nu$. Each standing wave may of course be thought of as two traveling waves, one going in each direction:

$$A \sin \omega \left(t - \frac{x}{c} \right) + A \sin \omega \left(t + \frac{x}{c} \right) = 2A \sin \omega t \cos \left(\omega \frac{x}{c} \right)$$

Thus $(kTl/c) d\nu$ thermal energy in the frequency range $d\nu$ is traveling to the right and an equal amount to the left. The amount of energy in this frequency range striking each end and being reflected per unit time is the above quantity divided by the time of transit $(c/l)^{-1}$ or $kT d\nu$.

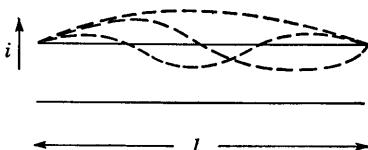
Assume now that the ends of the line are not shorted but are terminated by resistances matched to the line (see Fig. 6.6). These resistances are

¹ General fluctuation theory: H. B. Callen and T. A. Welton, *Phys. Rev.*, **83**, 34 (1951).

² Electrical fluctuations and noise: S. O. Rice, *Bell System Tech. J.*, **23**, 282 (1944); **24**, 47 (1945); R. Furth, *Proc. Roy. Soc. (London)*, **192**, 593 (1948).

³ Brownian motion: R. B. Barnes and S. Silverman, *Rev. Mod. Phys.*, **6**, 162 (1934); A. Ferguson, *Repts. Progr. in Phys.*, **2**, 1 (1935); M. C. Wang and G. E. Uhlenbeck, *Phys. Rev.*, **17**, 323 (1945); G. E. Uhlenbeck and L. S. Ornstein, *Phys. Rev.*, **36**, 832 (1930).

⁴ H. Nyquist, *Phys. Rev.*, **32**, 110 (1928).



dissipative elements so chosen that they absorb all the energy incident upon them from the line. The conclusions about the line itself are not critically dependent upon the terminations under the conditions assumed, and hence each resistance, which may be a function of ν , must absorb an energy kT per unit time per unit frequency interval. As these are considered as perfect thermal transformers, they must in equilibrium deliver energy to the line from their thermal surroundings at an exactly equal rate. Assuming that this electric energy appears in the line as if from many infinitesimal generators of electromotive force δe , in the frequency range $d\nu$ associated with each R_ν , the current of frequency in the range $d\nu$ will be $\delta i_\nu = \delta e_\nu / 2R_\nu$, and the power delivered from one R_ν to the other will be $\overline{(\delta i_\nu)^2 R_\nu} = \overline{(\delta e_\nu)^2} / 4R_\nu$. Equating this to $kT d\nu$, which is the power flowing along the line in each direction in the frequency range $d\nu$,

$$\overline{\delta e_\nu^2} = 4kT R_\nu d\nu \quad (6.14)$$

This form of the general relation between a thermally fluctuating quantity and a dissipative element is very convenient for electrical purposes. Stated in words it says that the mean-square fluctuating electromotive force of frequency between ν and $\nu + d\nu$ due to thermal oscillations is $4kT$ times the effective resistance at the frequency ν times the frequency interval $d\nu$. This clearly does not hold for frequencies so great that $h\nu$ is of the order of the thermal fluctuation energy kT , for there would be

few such energy quanta in the range kT and our statistics would not apply. Nevertheless, it is valid through the range of ν of electrical interest. δe_ν is frequently called the *noise voltage*. Taking T to be 300° abs and R a constant equal to 10^6 ohms, δe_ν is about 1.3μ volts in a 100-cycle frequency band. The total noise voltage evidently depends on the variation of R_ν with ν and the total frequency range accepted by one's detecting instrument.¹

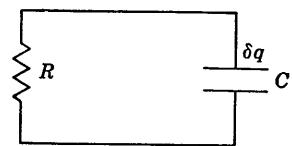


FIG. 6.7. Fluctuation of the charge on a condenser.

These matters may be looked at in another way which is convenient when the observations or measurements are not made in certain frequency bands but in a sequence of time intervals. Consider as an example the fluctuation of charge on a condenser of capacity C shunted by a resistance R . If the circuit of Fig. 6.7 contains a fluctuating thermal electromotive

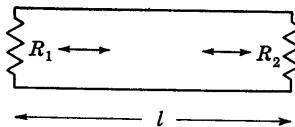


FIG. 6.6. Exchange of energy between terminating resistances by way of a parallel-wire line.

¹ N. H. Williams and W. S. Huxford, *Phys. Rev.*, **33**, 773 (1929).

force δe_T , the circuit equation is

$$\frac{\delta q}{C} + R \frac{d(\delta q)}{dt} = \delta e_T$$

If this equation is multiplied by δq , it becomes, since $\frac{d[(\delta q)^2]}{dt} = 2 \delta q \frac{d(\delta q)}{dt}$,

$$\frac{(\delta q)^2}{C} + \frac{1}{2} R \frac{d(\delta q)^2}{dt} = \delta e_T \delta q$$

If the terms are averaged, by integrating over an interval and dividing by that interval, the term on the right vanishes since δe_T is as often positive as negative and no correlation is assumed to exist between δe_T and δq . From the earlier discussion the average value of the fluctuating energy of a condenser is $\frac{1}{2}(\delta q)^2/C = \frac{1}{2}kT$. The averaged equation then yields

$$kT + \frac{R}{2} \frac{1}{(t_2 - t_1)} \overline{\int_{t_1}^{t_2} d(\delta q)^2} = 0$$

or

$$\overline{\Delta(\delta q)^2} = \overline{(\delta q)^2} = \frac{2kT}{R} \tau \quad (6.15)$$

where τ is the time interval separating measurements of $(\delta q)^2$. The same analysis would also apply, for instance, to the Brownian motion of a particle suspended in a fluid if δq is the displacement, R is the viscous damping force per unit velocity, and τ is the time interval between position measurements (Sec. 2.2).

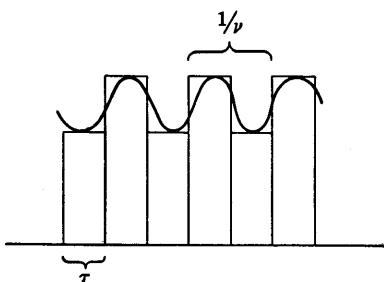


FIG. 6.8. Schematic illustration of the relation between time interval and associated frequency.

pulses spaced by the interval τ contributes principally to a Fourier component in the frequency spectrum of a period twice as long. In fact it may be shown rigorously that if a function $f(t)$ contains no frequencies outside the band from 0 to ν it is completely determined by giving its ordinates at a series of points spaced $1/2\nu$ sec apart.¹ Equations (6.14)

of a particle suspended in a fluid if δq is the displacement, R is the viscous damping force per unit velocity, and τ is the time interval between position measurements (Sec. 2.2).

Equations (6.14) and (6.15) are clearly very closely related. If Eq. (6.14) is written in terms of noise current for zero external resistance, $\delta i_r = \delta e_r / R_r$, then $(\delta i_r)^2 = 4kT \delta \nu / R_r$. The mean-square current from Eq. (6.15) is $(\delta q)^2 / \tau^2$ or $(\delta i)^2 = 2kT / R\tau$. As indicated in Fig. 6.8 a series of

¹ C. E. Shannon, Proc. IRE, 37, 10 (1949).

and (6.15) are essentially equivalent, the frequency band $\delta\nu$ being $1/2\tau$ sec $^{-1}$ in width.

An interesting result can be obtained from the application of statistical analysis to the fluctuations in the electron current between the cathode and anode of a diode. This current may be thought of as a stream of electrons striking the anode at the average rate n' . If the energy of interaction between the electrons constituting the current is small, i.e., if the current is not space-charge limited, the transit of each electron may be thought of as independent of all the others. The average number that will reach the anode in a time τ is then $\bar{n}'\tau$, and this is also equal to the average square fluctuation of this number if the current is not too small. Thus $(\delta\bar{n}'\tau)^2 = \bar{n}'\tau$ and, since the mean current \bar{i} is $\bar{n}'e$,

$$(\delta\bar{i})_{\tau}^2 = \frac{e\bar{i}}{\tau} \quad (6.16)$$

or using the relation $\delta\nu = 1/2\tau$ to convert this into the terms of a frequency interval,

$$\overline{(\delta i)}_{\nu}^2 = 2e\bar{i} d\nu \quad (6.16')$$

This is the so-called *shot-effect* equation, and the circuit of Fig. 6.9 may be used to determine the magnitude of the assumed elementary charge e of the electron. Williams and Huxford¹ made such measurements over a variation in \bar{i} of about fivefold, and the mean value of e that agrees with their observations is 1.58×10^{-19} coulomb, in excellent agreement with the accepted value of 1.60×10^{-19} coulomb for the electronic charge.

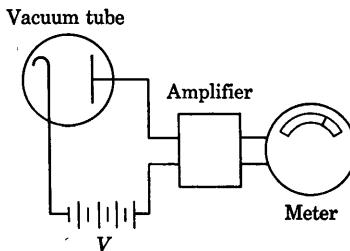


FIG. 6.9. Schematic of shot-effect circuit.

6.4. Quantum Statistics. The discussion in Sec. 6.2 was concerned with an ensemble of macroscopic systems having properties described by the basic classical mechanical and electrical laws governing such things as oscillatory circuits, weights and springs, etc. In this sense classical statistics is closely analogous to classical mechanics, which dealt adequately with the motion of large-scale objects. It is clear, however, that the results of Sec. 6.2 cannot be expected to apply to atoms or molecules and their components because some of the assumptions there made about systems are in contradiction with the properties that must be attributed to electrons, protons, neutrons, and photons. In the first place, an ensemble comprises a fixed but infinite number of systems whereas the actual fact is that a system with a finite number of atoms or

¹ N. H. Williams and W. S. Huxford, *Phys. Rev.*, **33**, 773 (1929).

molecules in any particular region is to be described. In low-energy processes of the type here dealt with, the number of electrons, protons, and neutrons remains fixed. However, the number of photons does not remain constant since photons can be emitted or absorbed and they appear or disappear from the physical assembly. More fundamentally, the systems of Sec. 6.2 were large-scale, recognizable, and identifiable entities whereas it has been seen that a basic property of elementary particles is their absolute identity in the sense that they are completely interchangeable and individually quite unidentifiable. One can only state, for instance, that an electron exists specified by certain parameters such as energy states in an atomic system or in a certain region of space with certain momentum components. This example brings up the final amendment that must be made to the assumptions of Sec. 6.2, namely, that the uncertainty principle implies a limitation upon the energy interval Δu that can be measured in a time Δt or the momentum interval Δp_x that may be used with a localization Δx , that is, $\Delta u \Delta t \cong \Delta p_x \Delta x \cong \hbar$. If these various quantum-mechanical concepts are incorporated in the statistical theory, it can be considered satisfactory, of course, only if the results continue to agree with those of Sec. 6.2 when applied to those large-scale phenomena for which classical statistics has proven to be satisfactory.

In considering atomic phenomena one can proceed exactly as in Sec. 6.2, with the understanding that the interval Δu is finite, up to the development of Eq. (6.5). That equation was based upon the hypothesis that one could meaningfully arrange the N systems in order and permute the subgroups of systems in particular energy intervals. Because of the Pauli principle such is not the case for atoms. One can only state how many atomic entities or elementary particles are described by a certain set of quantum numbers or by localization within certain ranges of position and momentum. To retain the description in terms of energy intervals as in Sec. 6.2, it will be assumed that the number of different sets of specifying parameters which correspond to an energy within the range u_j to $u_j + \delta u_j$ is g_j . This symbol thus represents the number of identifiable volumes or cells in a mathematical space with the chosen parameters, such as position coordinates and momenta, as coordinates. This combined coordinate and momentum space is known as a *phase space*. Evidently the volume that must be associated with the minimum element for the motion of an atom in three dimensions x , y , and z with momentum p_x , p_y , and p_z is $dx dy dz dp_x dp_y dp_z$ which by the uncertainty principle is of the order of \hbar^3 .

The number of quantum-mechanically distinguishable ways, then, in which n_j particles can each possess an energy between u_j and $u_j + \delta u_j$ is the number of ways of describing the content of particles in the g_j cells.

Reference to Fig. 6.10 shows that this is the same as the number of ways of arranging n_j particles which are indistinguishable and $(g_j - 1)$ internal cell walls which are indistinguishable, that is, $(n_j + g_j - 1)!/(g_j - 1)!n_j!$. The number of arrangements then of all the particles for all the δu_j energy intervals is

$$P_B = \prod_j \frac{(n_j + g_j - 1)!}{n_j!(g_j - 1)!} \quad (6.17B)$$

where \prod_j stands for the product of the factors for each energy interval δu_j .

This is proportional to the probability of occurrence of the particular occupancies n_j of the energy levels in terms of the physical parameters g_j and hence is the analogue of P in Eq. (6.5) for a classical distribution. The subscript B is used to identify the probability calculated with these particular assumptions, for they lead to a statistics known as the *Bose-Einstein statistics*.

These assumptions are satisfactory as far as the identity or indistinguishability property is concerned. However, elementary particles with intrinsic spin angular momenta $\frac{1}{2}\hbar$ have the additional characteristic that no two of them may be specified by the same total aggregate of

descriptive parameters such as quantum numbers or phase-space location. This means that if there are g_j separately specifiable cells in the energy range δu_j there can be at most g_j particles of this type in the energy range. If n_j is less than g_j , as illustrated in Fig. 6.11, there are a number of occupancy possibilities equal to the number of ways of choosing $(g_j - n_j)$ empty cells and n_j filled ones out of the total number of cells g_j , that is, $g_j!/(g_j - n_j)!n_j!$.

Such particles are said to obey the *Fermi-Dirac statistics*, and the total number of ways of distributing the particles over all the energy levels is

$$P_F = \prod_j \frac{g_j!}{(g_j - n_j)!n_j!} \quad (6.17F)$$

Recalling the discussion of Sec. 4.4 it is seen that there is a close relation between what may be called the parity of the angular momentum of a particle, i.e., whether its total angular momentum including spin is an even or odd multiple of $\hbar/2$, the symmetry of the wave function describing it, and the statistics appropriate for it. A consistent description of observations is achieved by associating antisymmetric wave functions

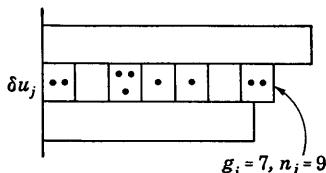


FIG. 6.10. Schematic of possible Bose-Einstein distribution.

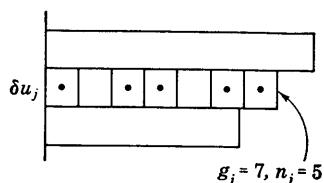


FIG. 6.11. Schematic of possible Fermi-Dirac distribution.

with odd integral multiples of $\hbar/2$ for the spin. Since these antisymmetric functions vanish if two particles have identical physical descriptions, the Fermi-Dirac statistics is appropriate for them. Symmetric wave functions are, however, to be employed for describing particles which have no spin or an integral multiple of \hbar units of angular momentum. Symmetric functions do not vanish if two particles of a system have the identical physical description; thus there is no limit upon the occupancy by them of phase-space cells, and the Bose-Einstein statistics is appropriate for them.

Equations (6.17B) or (6.17F) must be taken in place of Eq. (6.5) if one is to construct a statistics consistent with known properties of atoms or elementary particles. Equation (6.17F) applies if the spin of the particle is $n\hbar/2$, where n is an odd integer, and Eq. (6.17B) if the spin of the particle is $n\hbar$, where n is zero or an integer. In general there are of course the auxiliary conditions of Eqs. (6.4) representing constancy of total energy and total number of particles. In the case of photons the condition of total number of particles does not apply, but this can be included readily in the general case by setting the undetermined multiplier λ equal to zero for photons. The subsequent procedure is then exactly the same as in Sec. 6.2 except that the validity of Stirling's approximation is open to question for particular cases of energy intervals for which g_j or n_j are not large or for the Fermi-Dirac case if the cells are nearly all full and hence $g_j - n_j$ is not large. Neglecting these special limiting cases for the moment and applying Stirling's approximation,

$$\ln P_B = \Sigma[(n_j + g_j - 1) \ln (n_j + g_j - 1) - n_j \ln n_j - (g_j - 1) \ln (g_j - 1)]$$

$$\ln P_F = \Sigma[g_j \ln g_j - (g_j - n_j) \ln (g_j - n_j) - n_j \ln n_j]$$

and

$$\delta(\ln P_B) = \Sigma[\ln(n_j + g_j - 1) - \ln n_j] \delta n_j$$

$$\delta(\ln P_F) = \Sigma[\ln(g_j - n_j) - \ln n_j] \delta n_j$$

Multiplying the auxiliary conditions in the form $\delta N = 0 = \Sigma \delta n_j$ and $\delta U = 0 = \Sigma u_j \delta n_j$ by λ and μ , respectively, and adding to the above, with $\delta(\ln P)$ set equal to zero for the extreme (maximum) value,

$$\Sigma[\ln(n_j + g_j - 1) - \ln n_j + \lambda + \mu u_j] \delta n_j = 0 \quad \text{for } P_B$$

$$\Sigma[\ln(g_j - n_j) - \ln n_j + \lambda + \mu u_j] \delta n_j = 0 \quad \text{for } P_F$$

As in Sec. 6.2 the brackets must vanish to fulfill the condition, and therefore

$$n_j = \frac{g_j}{e^{-\lambda-\mu u_j} - 1} \quad (\text{Bose-Einstein}) \quad (6.18B)$$

$$n_j = \frac{g_j}{e^{-\lambda-\mu u_j} + 1} \quad (\text{Fermi-Dirac}) \quad (6.18F)$$

Unity has been neglected in comparison with g_j in the Bose-Einstein case which is legitimate if Stirling's approximation has been accepted. Equations (6.18B) and (6.18F) are the quantum-mechanical analogs of the classical equation (6.7). It is evident that the three equations are the same if unity is negligible in comparison with the exponential factor, and it will be seen that this corresponds to those situations in which classical statistics yields the correct result for atomic systems.

In the limiting case of very large energies, these equations approach $n_j = g_j e^{\lambda} e^{\mu u_j}$ if μ is negative. Under such conditions these equations must reduce to the classical one so that the constant μ here also must be identified with $-1/kT$. The constants λ and g_j must be determined by the particular physical situation being described. It is also possible to show that Eqs. (6.18) lead to a consistent thermodynamics for real atomic systems and to derive the partition function Z from them in terms of which the thermodynamic potentials ψ and Φ would be expressed. To do this, one may proceed on the hypothesis that Eq. (6.10') is the identification to be made between P and the thermodynamic function S . The logarithm of the Bose and Fermi probabilities may be written

$$\ln P = \Sigma[-n_j \ln n_j + ag_j \ln g_j - a(g_j - an_j) \ln (g_j - an_j)]$$

where $a = -1$ for the Bose statistics, and $a = +1$ for the Fermi statistics. Using Eqs. (6.18) in the form

$$n_j = g_j(e^{-\lambda} e^{-\mu u_j} + a)^{-1}, \quad g_j - an_j = g_j(1 + ae^{\lambda} e^{\mu u_j})^{-1}$$

and

$$\begin{aligned} \ln P &= \sum \left(n_j \ln \frac{g_j - an_j}{n_j} + ag_j \ln \frac{g_j}{g_j - an_j} \right) \\ &= \sum [n_j(-\lambda - \mu u_j) + ag_j \ln (1 + ae^{\lambda} e^{\mu u_j})] \\ &= -N\lambda - \mu U + a \sum g_j \ln (1 + ae^{\lambda} e^{\mu u_j}) \end{aligned}$$

or, as from Eq. (6.11) $\ln P = N(\ln Z - \mu U/N)$,

$$\ln Z = -\lambda + \frac{a}{N} \sum g_j \ln (1 + ae^{\lambda} e^{\mu u_j}) \quad (6.19)$$

This partition function satisfies Eqs. (6.9) as may be seen by substitution:

$$\begin{aligned} \frac{N}{\mu} \frac{\partial(\ln Z)}{\partial u_j} &= \frac{g_j}{e^{-\lambda} e^{-\mu u_j} + a} = n_j \\ \frac{\partial(\ln Z)}{\partial \mu} &= \frac{1}{N} \sum \frac{u_j g_j}{e^{-\lambda} e^{-\mu u_j} + a} = \frac{1}{N} \sum n_j u_j = \bar{u} \end{aligned}$$

Thus the entire description is consistent, and as in Sec. 6.2

$$\psi = -kT \ln Z, \quad \Phi = pV - kT \ln Z$$

These thermodynamic potentials per system are here the thermodynamic potentials per physical particle. For a physical ensemble containing N particles, ψ and Φ are N times as great. Finally it may be seen that the expression for Z [Eq. (6.19)] reduces to the classical expression except for a constant factor in the limit of large negative values of λ (Sec. 6.5).

6.5. Application of Quantum Statistics. *Black-body Radiation.* A problem of both practical importance and great historical interest is how the electromagnetic radiation within an enclosure whose walls are maintained at a temperature T can be described. Experimental observations on a system which approximates this can be made by examining the radiation that leaves such an enclosure by a hole that is so small in

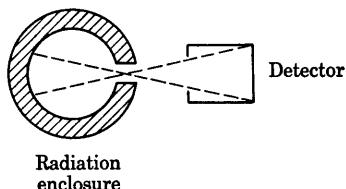


FIG. 6.12. Detection of radiation emerging from a black-body orifice.

comparison with the rest of the wall area that its disturbing effect can be ignored. Then the rate at which radiant energy will leave the enclosure and enter the detector, assuming the radiation within the enclosure is isotropic, is the product of the energy density of the radiation in the enclosure, its velocity, and the solid angle subtended at the enclosure hole by the detector opening (see Fig. 6.12). If a dis-

persing instrument is interposed between source and detector, the dependence of energy density upon frequency or wavelength can be determined also. The energy density of the radiation within the enclosure and its spectral distribution may be calculated in order to compare the theoretical prediction with experiment.

The applicability of Eq. (6.18B) will be assumed, i.e., photons obey the Bose-Einstein statistics. The consequences of this assumption are found to be in agreement with experiment, whereas the consequences of the assumption that photons obey the Fermi-Dirac statistics are not. Since there is no condition for the constancy of the number of photons in the statistical assembly within the enclosure, it is evident that the constant λ need not appear. Setting it equal to zero,

$$n_j = \frac{g_j}{e^{u_j/kT} - 1}$$

Since the energy of a photon is $h\nu$, ν provides the energy scale for u . The time of observation for this experiment is very great compared to $1/\nu$; for the frequency range under consideration, the uncertainty principle $\delta\nu \delta t \cong 1$ imposes a negligible restriction upon $\delta\nu$. So n_j , the num-

ber of photons in the energy range δu_j , is essentially $n_\nu d\nu$, where n_ν is the density of photons per unit frequency range and $d\nu$ is essentially an infinitesimal element of frequency range. It only remains to determine g_j , the number of cells in the phase space corresponding to the energy interval δu_j .

This may be done by considering the number of standing-wave patterns in the enclosure as a function of frequency and translating this into energy terms by the relation $u = h\nu$. It may be shown that the distribution function is not sensitive to the particular shape of the enclosure and a cubic box of side l (Fig. 6.13) is the simplest geometry for analysis. Two waves of equal amplitude and angular frequency ω in the direction $+k$ and $-k$ give rise to a standing wave:

$$e^{i(\omega t + \mathbf{k} \cdot \mathbf{r})} + e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} = 2e^{i\omega t} \cos \mathbf{k} \cdot \mathbf{r}$$

In a cubic box of uniform walls, $\cos \mathbf{k} \cdot \mathbf{r}$ must have the same value over the sides at $x = 0$ and $x = l$, that is,

$$k_x \cdot 0 + k_y y + k_z z = k_x \cdot l + k_y y + k_z z \pm 2\pi n_x$$

and similarly for y and z , so

$$k^2 = \frac{4\pi^2}{\lambda^2} = k_x^2 + k_y^2 + k_z^2 = \frac{4\pi^2}{l^2} (n_x^2 + n_y^2 + n_z^2)$$

or $n_x^2 + n_y^2 + n_z^2 = (l\nu/c)^2$, which represents the limitation on the n 's for the frequency ν . If the n 's are considered to represent the units along three perpendicular axes, this is the equation of a sphere of radius $l\nu/c$ so the number of unit volumes in this sphere, $\frac{4}{3}\pi(l\nu/c)^3$, is the number of sets of values of n_x , n_y , n_z for which standing waves exist in the box with frequencies less than ν . If this number is multiplied by 2 for the two possible polarizations of the electromagnetic wave, one has the number of describable waves of frequency less than ν . The number of photons of energy between u and $u + du$ is

$$d \left(\frac{8\pi}{3} \frac{l^3}{c^3} \frac{u^3}{h^3} \right) = g_j = \frac{8\pi V u^2 du}{(ch)^3}$$

where V is the volume of the enclosure. Finally, writing $n_j = n_\nu d\nu$ and $u = h\nu$,

$$n_\nu d\nu = \frac{8\pi V}{c^3} \frac{\nu^2 d\nu}{e^{h\nu/kT} - 1}$$

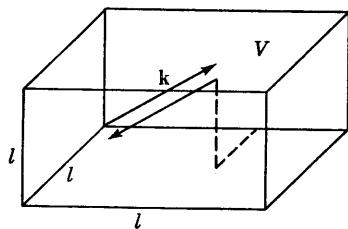


FIG. 6.13. Construction for analyzing normal modes of enclosed radiation.

and the energy u'_ν per unit volume per unit frequency range is $\hbar\nu n_\nu/V$, so

$$u'_\nu d\nu = \frac{8\pi\hbar\nu^3}{c^3} \frac{d\nu}{e^{\hbar\nu/kT} - 1} \quad (6.20)$$

This is *Planck's black-body radiation law*, and it was the agreement between the predictions of this law and experiment that provided the first evidence for the quantum theory of radiation. It should be pointed out that g_i could be obtained by considering the number of phase-space cells $dp_x dp_y dp_z dx dy dz$ corresponding to all orientations of the photon

propagation vector when the photon is confined within the volume V . The phase-space volume (illustrated in Fig. 6.14) corresponding to p being between p and $p + dp$ in the volume V is $4\pi p^2 dp V$ or, allowing for twice this volume for the two directions of polarization, $8\pi p^2 V dp$. As p for a photon is $\hbar\nu/c$, this is $8\pi\hbar^3\nu^2 V d\nu/c^3$. For this to agree with the previous result the elementary volume of phase space must equal h^3 , i.e.,

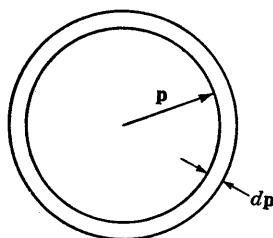


FIG. 6.14. Momentum-space element.

$$dp_x dp_y dp_z dx dy dz = h^3 \quad (6.21)$$

The uncertainty principle suggests this relation, and the agreement of Eq. (6.20) with experiment therefore determines the smallest elementary physical volume of phase space necessary to describe a photon wave vector.

The correctness of Eq. (6.20) has been demonstrated by Coblenz and many other investigators,¹ and the nature of the dependence of u'_ν on $\hbar\nu/kT$ is shown in Fig. 6.15. If $x = \hbar\nu/kT$, then

$$u'_\nu = 8\pi h \left(\frac{kT}{hc} \right)^3 x^3 (e^x - 1)^{-1}.$$

This is a maximum for $x = 2.8215$. If the Fermi-Dirac statistics had been assumed for photons, u'_ν would have been proportional to

$$x^3(e^x + 1)^{-1}$$

and the maximum would have occurred for $x = 3.131$. The experimental results are quite accurate enough to decide in favor of the Bose-Einstein statistics for photons. In such experiments the wavelength is generally measured rather than the frequency, and since $\nu = c/\lambda$

$$u'_\lambda d\lambda = \frac{8\pi hc}{\lambda^5} \frac{d\lambda}{e^{hc/\lambda kT} - 1}$$

¹ R. Birge, *Rev. Mod. Phys.*, **1**, 56 (1929); American Institute of Physics, "Temperature Symposium," Reinhold Publishing Corporation, New York, 1941.

or

$$u'_\lambda = 8\pi ch \left(\frac{kT}{hc} \right)^5 x^5 (e^x - 1)^{-1}$$

This has a maximum value for $x = 4.9651$, giving Wien's displacement law

$$(\lambda T)_{\max} = \frac{hc}{4.9651k}$$

Thus the product of the wavelength of maximum radiation intensity and the enclosure temperature is a constant involving h , c , and k . From

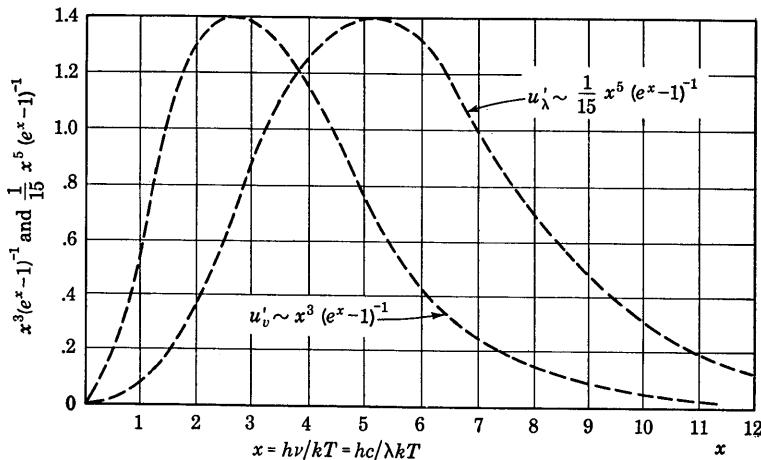


FIG. 6.15. Black-body radiation energy-distribution curves.

measurements of $(\lambda T)_{\max}$, hc/k is found to be 1.436 ± 0.002 cm-deg. From the values of the constants h , c , and k the value of hc/k is calculated to be 1.4387 with an accuracy of 50 parts per million. This agreement is very good considering the difficulty of the experiment. The total energy density may be obtained by integrating u'_v or u'_λ giving the Stefan-Boltzmann law (see Prob. 6.13) which has also been satisfactorily checked by experiment.

It is also of interest to consider this radiation or assembly of photons to be in equilibrium with the atoms of the walls of the container. Certain general conclusions concerning the radiation process may then be derived without any particular assumptions (Sec. 3.6) regarding the quantum-mechanical details of the radiation process. Let it be assumed that there is an atomic process whereby an atom which is in an excited state spontaneously emits an energy quantum $h\nu$ changing from an energy state designated by a to one designated by b , $E_a - E_b = h\nu$. The probability of this process occurring per unit time per atom in state a is written

$A(ab)$ or A_{ν} . Thus the rate of change of the number of atoms in state a by radiation processes to lower states b is

$$\frac{\partial N(a)}{\partial t} = - \sum_b A(ab)N(a)$$

If this process alone were operative the number of atoms in state a would decrease exponentially in accordance with the equation $N(a) = N_0(a)e^{-t/\tau}$, where $\tau = \left[\sum_b A(ab) \right]^{-1}$.

If the atoms are in equilibrium with thermal radiation in an enclosure, the effects of this radiation must be taken into account as both enabling the reverse processes to occur whereby atoms absorb energy and are returned from the various b states to state a and as possibly stimulating a greater rate of change from state a to the other states by the emission of radiation. It is assumed that the effect of the radiation on these processes is proportional to the energy density of the radiation at the effective frequency ν . In accordance with this assumption the rate at which atoms are raised to state a by the radiation is

$$\sum_b N(b)B'(ba)u'_{\nu}(ab)$$

where $B'(ba)$ is the probability of an atom absorbing a quantum of frequency $\nu(ab)$ per unit time and $u'_{\nu}(ab)$ is the energy density at that frequency. Similarly, the rate at which induced emission depletes state a is

$$\sum_b N(a)B(ab)u'_{\nu}(ab)$$

where $B(ab)$ is the probability of induced emission per unit time per unit energy density of the ambient radiation of the proper frequency. Equilibrium then requires the equality of the rates of accretion and depletion of level a , or

$$\sum_b N(b)B'(ba)u'_{\nu}(ab) = \sum_b N(a)[A(ab) + B(ab)u'_{\nu}(ab)]$$

The argument, which was initially proposed by Einstein, then proceeds upon the basis of a general principle called the *principle of detailed balancing* in accordance with which the rate of each elementary process is equal to the rate of its precise inverse in a situation of statistical equilibrium. That is, the rate of transfer of atoms from any particular level b to a is the same as that from a to the same b , as indicated in Fig. 6.16. This principle is justified by the agreement between the

resulting theory and experimental observation. In accordance with this principle the summation sign may be dropped and the preceding equation is true without it, i.e.,

$$N(b)B'(ba)u'_\nu(ab) = N(a)[A(ab) + B(ab)u'_\nu(ab)]$$

Then one may apply the laws of classical statistics to the number of the localized atoms in the enclosure walls in the assumed energy states E_a

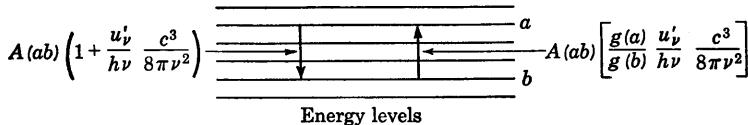


FIG. 6.16. Einstein coefficients.

and E_b ; allowing the possibility that the statistical weight g of the states a and b may be different:

$$N(a) = g(a)e^{-E_a/kT}, \quad N(b) = g(b)e^{-E_b/kT}$$

Writing these in the preceding equation and solving for $u'_\nu(ab)$,

$$u'_\nu(ab) = \frac{A(ab)/B(ab)}{[g(b)B'(ba)/g(a)B(ab)]e^{h\nu/kT} - 1}$$

Comparing this with Eq. (6.20) it is seen that for agreement with experiment the following relations must hold between the Einstein coefficients:

$$\begin{aligned} g(b)B'(ba) &= g(a)B(ab) \\ A(ab) &= B(ab) \frac{8\pi h\nu^3}{c^3} \end{aligned} \tag{6.22}$$

The first of these relations is very interesting in showing the equality between the probability of induced emission and absorption for energy states of equal weighting. The probability of interaction of radiation with an atom either in emission or absorption is proportional to the weighting factor or the number of quantum-mechanically describable configurations corresponding to the particular energy state. These matters will be seen to have a bearing on the discussion in Sec. 3.3. Utilizing the second of the preceding relations, the probability of emission can be written

$$A(ab) + B(ab)u'_\nu(ab) = A(ab) \left[1 + u'_\nu(ab) \left(\frac{8\pi h\nu^3}{c^3} \right)^{-1} \right]$$

The second term in the brackets is the quantum energy density per unit volume per unit frequency interval ($u'_\nu/h\nu$) divided by the weighting factor or number of phase-space cells ($8\pi\nu^2/c^3$) per unit frequency range

at the frequency ν . Thus the induced emission and also the induced absorption processes are proportional to the phase-space density of appropriate quanta of radiant energy.

Gas Laws. The problem of a gas is that of some number of atoms, say N , in an enclosure of volume V , the atoms being free to occupy the entire volume whatever size it may be. In order that Eqs. (6.18) may be applied, g_j and λ must be determined for this situation. The quantity g_j depends, of course, upon the way in which the energy of a molecule can be

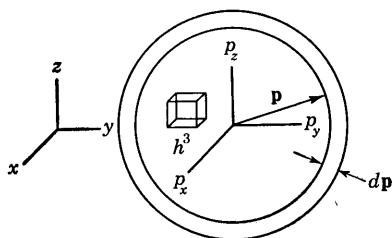


FIG. 6.17. h^3 unit element in six-dimensional coordinate and momentum (phase) space.

described, for it is the number of cells in the phase space corresponding to δu_j . The simplest case which will illustrate the method is that of a monoatomic gas for which there is no energy of rotation or vibration and for which there is no mutual potential energy between the molecules. Thus the energy is simply the kinetic energy of translation of the atoms. The influence of the gravitational field of the earth is neglected; hence the density is uniform. It is further assumed that there are no electric or magnetic fields to affect the atoms.

Then much as in the case of photons the number of phase-space cells corresponding to momentum components between p and $p + dp$ is (see Fig. 6.17)

$$\frac{4\pi p^2 dp}{dp_x dp_y dp_z} = \frac{4\pi p^2 dp}{h^3} dx dy dz$$

as $dp_x dp_y dp_z dx dy dz = h^3$. The energy is independent of x , y , and z , so these coordinates may be integrated over directly to yield for a volume V

$$g_j = \frac{4\pi p^2 V dp}{h^3}$$

If the atoms have certain internal structure upon which the distribution function may depend, there will in general be other contributions to the factor g . For instance, the atoms may have angular momenta including spin, and if the value of this is $\sqrt{J(J+1)}\hbar$ it can have $(2J+1)$ distinguishable orientations (Sec. 3.4). This factor, just as the factor 2 for photon polarization, should be included as a weighting factor in g_j and will be written in general as w in the following discussion. Then since $\sum n_j = \int_0^\infty n(p) dp = N$,

$$N = \int_0^\infty \frac{4\pi p^2 w V dp / h^3}{e^{-\lambda e^{-\mu u}} + a} \quad (6.23)$$

where $a = -1$ for the Bose statistics, and $a = +1$ for the Fermi-Dirac statistics. This equation determines λ , the only constant as yet unknown. When λ is determined, $\bar{u} = \frac{1}{N} \sum u_j n_j$ can be immediately calculated, and this result contains the answer to most questions regarding the behavior of gas atoms in thermal equilibrium.

Writing $n' = N/V$, the number of atoms per unit volume, and

$$x = p(2mkT)^{-\frac{1}{2}}$$

to simplify the variable of integration, Eq. (6.23) becomes

$$n' = 4\pi w \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \int_0^\infty \frac{x^2 dx}{e^{-\lambda e^{x^2}} + a} \quad (6.24)$$

The mean kinetic energy of translation \bar{u} , which is $\int \frac{p^2}{2m} n(p) dp$ divided by N , is

$$\bar{u} = 4\pi w \frac{kT}{n'} \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \int_0^\infty \frac{x^4 dx}{e^{-\lambda e^{x^2}} + a} \quad (6.25)$$

The actual determination of λ requires the performance of the integration indicated in Eq. (6.24), but this cannot in general be done in a simple closed fashion. However, without any particular assumptions, Eq. (6.19) permits the determination of certain quantities of thermodynamic interest. Using the value of g_i previously determined, Eq. (6.19) becomes

$$\ln Z = -\lambda + \frac{4\pi a}{N} wV \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \int_0^\infty x^2 \ln (1 + ae^{\lambda e^{-x^2}}) dx \quad (6.26)$$

Then, since $\psi = \bar{u} - ST = -kT \ln Z$ and $d\bar{u} = T dS - p dV$ from the first law, $p = -(\partial\psi/\partial V)_T$ or

$$\begin{aligned} p &= kT \frac{\partial}{\partial V} (\ln Z)_T \\ &= \frac{4\pi a k T w}{N} \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \int_0^\infty x^2 \ln (1 + ae^{\lambda e^{-x^2}}) dx \end{aligned} \quad (6.27)$$

If this is integrated by parts using the factors $x^2 dx$ and $\ln (1 + ae^{\lambda e^{-x^2}})$,

$$p = 4\pi w \frac{kT}{N} \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \int_0^\infty \frac{2}{3} \frac{x^4 dx}{e^{-\lambda e^{x^2}} + a} = \frac{2}{3} \frac{\bar{u}}{V}$$

Thus for any value of λ and either type of statistics, one has the general result for a simple monatomic gas that

$$pV = \frac{2}{3}\bar{u} \quad (6.28)$$

Also, of course, on integrating Eq. (6.26) by parts,

$$\ln Z = -\lambda + \frac{2}{3} \frac{\bar{u}}{kT}$$

$$\psi = \lambda kT - \frac{2}{3} \bar{u}$$

and

$$S = \frac{\bar{u} - \psi}{T} = -\lambda k + \frac{5}{3} \frac{\bar{u}}{T} \quad (6.29)$$

but without a knowledge of λ these expressions cannot be evaluated in terms of the usual thermodynamic variables.

To proceed further it is thus necessary to make assumptions regarding λ and see to what physical situations these assumptions correspond. If λ is assumed to be very large and negative so that a may be neglected over almost the entire range of x , Eq. (6.24) becomes

$$n' \cong 4\pi w \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} e^\lambda \int_0^\infty x^2 e^{-x^2} dx$$

$$\cong 4\pi w \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \frac{e^\lambda}{2} \int_0^\infty e^{-x^2} dx$$

on integrating by parts. The last integral is a well-known one equal to $\sqrt{\pi}/2$, so

$$e^{-\lambda} = \frac{w}{n'} \left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} \quad (6.30)$$

Under the present assumption (λ large and negative) $e^{-\lambda}$ must be large. Inserting the values of k and h and writing m in terms of molecular-weight units and n' in terms of standard gas conditions (6.025×10^{23} atoms in 22.4 liters is unit density, $\rho = 1$),

$$e^{-\lambda} \cong 6.9 \frac{(MT)^{\frac{3}{2}}}{\rho} w$$

If M and T are large and ρ is small, the approximation is good. For instance, helium ($M = 4$, $w = 1$) at normal temperature and density yields $e^{-\lambda} \approx 3 \times 10^5$, and the approximation is seen to be quite justified for ordinary ranges of temperature and pressure. Under these circumstances, from Eq. (6.27),

$$p = \frac{4\pi akT}{N} \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} w \int x^2 a e^\lambda e^{-x^2} dx$$

$$= n' \frac{kT}{N} = \frac{kT}{V} \quad (6.31)$$

where Eq. (6.30), $\int_0^\infty e^{-x^2} dx = \sqrt{\pi}/2$, and the approximation for $\delta \ll 1$, $\ln(1 + \delta) \cong \delta$ have been used. This is the perfect-gas law for a single particle, and for N particles it would be $pV = NkT$. Thus the theory leads to the perfect-gas law in the limiting case, and from Eqs. (6.28) and (6.29)

$$S = k \ln \left[\frac{w}{n'} \left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} \right] + \frac{5}{2}k$$

The entropy is commonly written for a mole of N molecules where $Nk = R$, in which case

$$\begin{aligned} S &= R \left(\frac{5}{2} + \frac{3}{2} \ln \frac{2\pi mkT}{h^2} - \ln \frac{n'}{w} \right) \\ \Psi &= N\Psi = RT \left(\ln \frac{n'}{w} - 1 - \frac{3}{2} \ln \frac{2\pi mkT}{h^2} \right) \end{aligned} \quad (6.32)$$

One could explore the first approximations to the above results for values of λ that are not so large and negative, i.e., for lower temperatures and higher densities, but this is unprofitable because the small departures from Eqs. (6.31) and (6.32) would not be directly comparable with experiment. Interatomic forces which have here been neglected cause larger experimental deviations from the perfect-gas laws than does the departure from classical statistics. But at the opposite extreme for which λ is quite small (it is zero for radiation as has been seen), important new situations arise which lead to nonclassical results as will be seen in the following discussion and in Secs. 7.5 and 7.7.

Thomas-Fermi Atom. It is interesting to see to what extent the foregoing ideas can be applied to the electron structure of an atom if it contains a reasonably large number of electrons, say 10 or more. Though the result obtained is of an approximate character it provides a simple, general, and frequently very useful approach to the properties of heavy polyelectronic atoms. In the approximation here presented this example is more an illustration of the uncertainty principle than of quantum statistics since it is assumed that all the electrons occupy the lowest available energy states possible within the limitations imposed by the Pauli principle. This is equivalent to assuming that the temperature is very low and there is no tendency for the electrons by reason of excess energy to occupy other than those states for which the total energy is the least. For simplicity the effects of electron exchange are here neglected, though a modification to take these into account has been proposed by Dirac.¹ Under these simplest circumstances the number of electrons in the energy state u_j is g_j , i.e., $n_j = g_j$. In terms of the number $n_p dp$ of

¹ P. A. M. Dirac, *Proc. Cambridge Phil. Soc.*, **26**, 376 (1930).

electrons with momenta between p and $p + dp$,

$$n_p dp = 2 \frac{4\pi p^2 dp}{h^3} d\tau$$

where the factor 2 comes from the possible spin orientations and $d\tau$ is an elementary unit of spatial volume. Then the number of electrons of all momenta per element of volume is

$$n' = \int_0^{p_{\max}} n_p dp = \frac{8}{3} \frac{\pi}{h^3} p_{\max}^3$$

where p_{\max} is the upper limit of the momentum necessary to accommodate all the electrons.

In order that the most energetic electrons should not escape from the atomic system, the energy corresponding to p_{\max} , $p_{\max}^2/2m_0$, must be less than the electrostatic binding energy of an electron in the potential field of the atomic system ($-e\varphi$). Thus $p_{\max}^2 = -2m_0 e\varphi$ is the limiting condition, and the equation for n' in terms of φ becomes

$$n' = \frac{8\pi}{3h^3} (-2m_0 e\varphi)^{\frac{3}{2}}$$

A second relation exists between n' and φ because of Poisson's equation:

$$\nabla^2 \varphi = - \frac{\rho}{\epsilon_0} = - \frac{n' e}{\epsilon_0}$$

where ρ is the charge density, en' . Eliminating n' ,

$$\nabla^2 \varphi = - \frac{8\pi e}{3h^3 \epsilon_0} (-2m_0 e\varphi)^{\frac{3}{2}}$$

The atomic potential φ may be taken as spherically symmetrical, i.e.,

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right)$$

and using the substitution $\varphi(r) = (-Ze/4\pi\epsilon_0 r)X(r)$,

$$\frac{d^2 X}{dr^2} = \frac{8\pi e^3}{3h^3 \epsilon_0} (2m_0)^{\frac{3}{2}} \left(\frac{Z}{4\pi\epsilon_0} \right)^{\frac{1}{2}} r^{-\frac{1}{2}} X^{\frac{3}{2}}$$

This may be further simplified by the substitution $r = bx$, where

$$b = \left(\frac{9}{128\pi Z} \right)^{\frac{1}{3}} \frac{\epsilon_0 h^2}{m_0 e^2} = \frac{\pi}{4} \left(\frac{9}{2\pi} \right)^{\frac{1}{3}} Z^{-\frac{1}{3}} a_0 = 4.681 \times 10^{-11} Z^{-\frac{1}{3}} \quad \text{m}$$

and a_0 is the Bohr-orbit constant 0.529×10^{-10} m of Sec. 3.4. The equation between the dimensionless quantities X and x is then

$$x^{\frac{1}{3}} \frac{d^2 X}{dx^2} = X^{\frac{1}{3}} \quad (6.33)$$

This differential equation is unusual among those commonly occurring in atomic physics in being nonlinear; the sum of separate solutions is not in general a solution also. A general solution in a closed form is not known; but X may be expanded in a semiconvergent power series of $x^{\frac{1}{3}}$ about the origin, and numerical integration yields X in terms of the assumed boundary conditions. One general condition is that as r approaches zero the potential must approach the Coulomb potential of the nucleus with the charge Ze , i.e., for r or x small, $\varphi \rightarrow -Ze/4\pi\epsilon_0 r$ and hence $X \rightarrow 1$. This determines one of the constants of integration, and by substitution of the power series in Eq. (6.33) the series for X may be shown to be

$$X = 1 + s_0 x + \frac{4}{3} s_0^{\frac{1}{3}} x^{\frac{1}{3}} + \frac{2}{5} s_0 x^{\frac{2}{3}} + \frac{1}{3} s_0^{\frac{2}{3}} x^{\frac{2}{3}} + \frac{3}{70} s_0^{\frac{2}{3}} x^{\frac{4}{3}} + \dots$$

where $-s_0$ is the slope of the curve at the origin. The outer boundary condition may be chosen to correspond to a neutral atom or an ion by setting the field at the atomic boundary equal either to zero or to $ne/4\pi\epsilon_0 r^2$, where ne is the net ionic charge. In the case of a neutral atom the condition at the boundary is $-\partial\varphi/\partial r = 0$ or $\partial X/\partial x = X/x$. This occurs at finite values of x if s_0 is less than -1.588 . If s_0 has this particular value, the curve $X(x)$ approaches the x axis asymptotically and the solution corresponds to an isolated neutral atom. A complete discussion of the solution for various cases and tables of numerical values of X are to be found in the references.¹

Figure 6.18 is a plot of X as a function of x for an isolated neutral atom. The quantities of particular physical interest are immediately derivable from such a graph. These universal equations are applicable to all heavy atoms on this theory.

$$\begin{aligned} \varphi &= \frac{-Ze}{4\pi\epsilon_0 r} X = \frac{-Ze}{4\pi\epsilon_0 b} \frac{X}{x} = -30.75 Z^{\frac{1}{3}} \frac{X}{x} \quad \text{volts} \\ E_r &= -\frac{\partial\varphi}{\partial r} = -\frac{1}{b} \frac{\partial\varphi}{\partial x} = 6.570 \times 10^{11} Z^{\frac{1}{3}} \frac{\partial}{\partial x} \left(\frac{X}{x} \right) \quad \text{volts m}^{-1} \\ \rho &= n'e = \frac{8\pi e}{3} \left(\frac{Z}{2\pi^2 b a_0} \right)^{\frac{1}{3}} \left(\frac{X}{x} \right)^{\frac{1}{3}} = -1.243 \times 10^{11} Z^2 \left(\frac{X}{x} \right)^{\frac{1}{3}} \quad \text{coulombs m}^{-3} \\ \rho' &= 4\pi r^2 \rho = 4\pi b^2 x^2 \rho = -3.422 \times 10^{-9} Z^{\frac{1}{3}} x^{\frac{1}{3}} X^{\frac{1}{3}} \quad \text{coulombs m}^{-1} \end{aligned}$$

¹ W. Bush and S. H. Caldwell, *Phys. Rev.*, **38**, 1898 (1931); L. Brillouin, *L'Atome de Thomas-Fermi*, Hermann & Cie, Paris, 1934; J. C. Slater and H. M. Krutter, *Phys. Rev.*, **47**, 559 (1935); R. P. Feynman, N. Metropolis, and E. Teller, *Phys. Rev.*, **75**, 1561 (1949).

Here ρ' is the radial charge density. ϕ , E_r , and ρ all become infinite for $x = 0$, but the charge per unit radius ρ' does not;

$$\frac{\rho'}{3.422 \times 10^{-9} Z^{\frac{1}{3}}} = x^{\frac{1}{3}} X^{\frac{1}{3}}$$

which is a universal curve for all atoms, is plotted in Fig. 6.19. This theory approximates the actual distribution of electrons but of course

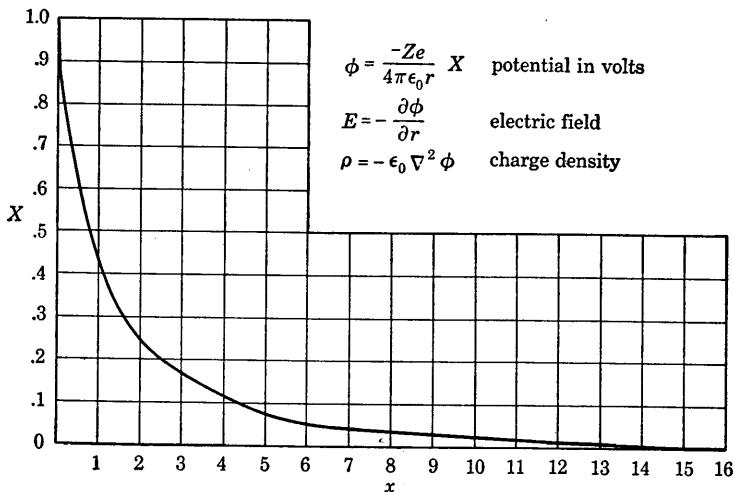


FIG. 6.18. Thomas-Fermi function.

does not show the fine detail of the shell structure that can be obtained only by more refined methods. It provides a very useful approximation for scattering calculations involving heavy atoms, for it is evident that

$U(r)$ of Sec. 5.1 is the same as ρ'/e . The atomic scattering factor F of Sec.

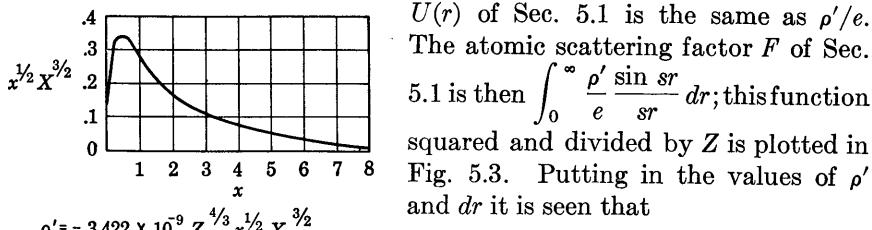


FIG. 6.19. Thomas-Fermi radial charge density.

$$f(sb) = \frac{F(sb)}{Z} = \int_0^\infty x^{\frac{1}{3}} X^{\frac{1}{3}} \frac{\sin sbx}{sbx} dx$$

is a universal function for all atoms provided that the quantity

$$sb = \left[\frac{4\pi}{\lambda} \sin \frac{\phi}{2} \right] \left[\frac{\pi}{4} \left(\frac{9}{2\pi} \right)^{\frac{1}{3}} \frac{1}{Z^{\frac{1}{3}}} a_0 \right]$$

$$= \frac{11.12 a_0}{\lambda Z^{\frac{1}{3}}} \sin \frac{\phi}{2}$$

is the same, i.e., for large Z a smaller λ must be used to obtain the same scattering pattern. If the wavelength is proportional to the linear atomic dimensions, the scattering pattern is the same. $f(s)$ is plotted in Fig. 5.4.

From the fact that $r = bx$, where b varies as $Z^{-\frac{1}{3}}$, it is evident that the radius vector out to any representative point of an electron structure is proportional to $Z^{-\frac{1}{3}}$, or roughly speaking the radius of an atom is proportional to $Z^{-\frac{1}{3}}$; heavy atoms are smaller in radius by this factor. Finally, it may be noted that $X = ax^n$ is a solution of Eq. (6.33) if $a = 144$ and $n = -3$. This solution is unsatisfactory for small values of x : for small x , X does not approach unity, but it is satisfactory for the boundary condition $x \rightarrow \infty$ and gives the behavior of the curve of Fig. 6.18 at large distances from the center of the atom. From this approximation the potential in volts may be shown to be given by $4.42 \times 10^{-3}(r/a_0)^{-4}$ for values of r large compared to a_0 for any heavy atom. For further applications of the Fermi-Thomas theory to atomic systems the references should be consulted.

PROBLEMS

6.1. In many interesting cases of statistical analysis, m , $m - n$, and n of Eq. (6.1) are not all large. Such would, for instance, be the case in the counting example of the text if the counts occurred so infrequently that p was very small and the number of counts occurring on the average in τ , $\bar{n} = mp = \tau p/\delta t$, was not large. Assuming that $m \gg 1$ and $p = \bar{n}/m \ll 1$, and using the definition of the exponential function

$$\left(1 - \frac{a}{x}\right)^x \underset{x \rightarrow \infty}{=} e^{-a}$$

show that Eq. (6.1) becomes

$$\begin{aligned} P_n &= \frac{m!}{(m-n)!n!} \left(\frac{\bar{n}}{m}\right)^n \left(1 - \frac{\bar{n}}{m}\right)^{(m-n)} \\ &\cong \frac{\bar{n}^n e^{-\bar{n}}}{n!} \end{aligned}$$

This is known as the *Poisson distribution law*. See Fig. 6.2.

6.2. Background counts as registered by a given counter system are measured for a time t_1 during which the number n_1 is recorded. A source of radiation is then placed near the counter, and a number of counts n_2 is recorded in a time t_2 . The counting rate due to the source alone is $n_2/t_2 - n_1/t_1$. Taking the standard deviation of a sum or difference as the square root of the sum of the squares of the separate deviations, and assuming that the two measurements are made for such a length of time that the deviations in n_1 and n_2 are the same, show that the total time that must be taken for these measurements to achieve an acceptable fractional precision f is

$$t_{\text{total}} = 2 \frac{1 + 2n'_b/n'_s}{n'_s f^2}$$

where n'_b and n'_s are the background and source rates, respectively.

6.3. A counter is set up to measure the relative intensity of the radiation from two sources of similar character but unequal strengths. The first series of background readings is taken when neither source is present. This series shows that in 85 of the 1-min intervals no counts were registered, in 166 of the 1-min intervals one count was registered, in 172 of the 1-min intervals two counts were registered, etc. Then source *A* was placed in position, and the number of 1-min intervals with the corresponding number of counts obtained in them are given in the next series. Then source *A* was replaced by source *B* and the final series was obtained. Calculate the average counting rate in each case and the precision (\sqrt{n}) with which it may be assumed to be known. Calculate the deviation σ for the first and second series. Plot the number of intervals containing x counts as a function of x , and compare the points with the curve obtained from the Poisson formula of Prob. 6.1 for the first and second series; compare the second and third series with the Laplace formula, Eq. (6.3). After subtracting the background counts (using the relation that the deviation of a sum is the square root of the sum of the squares of the separate deviations), show that source *B* is five times as strong as source *A* with an accuracy of about 1.3 per cent.

Series 1 (Background)

Number of intervals:	85	166	172	113	55	21	8	2	1
Number of counts:	0	1	2	3	4	5	6	7	8

Series 2 (Source A)

Number of intervals:	0	1	5	20	48	96	160	221	287	322	320
Number of counts:	0	1	2	3	4	5	6	7	8	9	10

Number of intervals:	291	242	180	133	89	55	33	17	10	4	2
Number of counts:	11	12	13	14	15	16	17	18	19	20	21

Series 3 (Source B)

Number of intervals:	1	1	3	4	6	10	12	16	17	20	24
Number of counts:	27	28	29	30	31	32	33	34	35	36	37

Number of intervals:	26	29	30	30	32	32	30	29	26	23	21
Number of counts:	38	39	40	41	42	43	44	45	46	47	48

Number of intervals:	17	15	11	10	7	3	3	1	1
Number of counts:	49	50	51	52	53	54	55	56	57

6.4. Two counters are so arranged in a circuit as to register a coincidence if a discharge in one counter occurs in the small interval Δt before or Δt after the discharge in the other counter. Assuming that the counters are recording two quite different series of events, so that any coincidences are quite by chance, at the average rates \bar{n}'_1 and \bar{n}'_2 , respectively, show that the chance that there will be a single coincidence per unit time is approximately $2 \Delta t \bar{n}'_1 \bar{n}'_2$ if this quantity is very small. A group of *S* counters are so arranged in a circuit that a coincidence of all *S* of them is recorded if they all discharge within a time of length $S \Delta t$. Show that the probability of a chance coincidence per unit time of all *S* counters is $(S \Delta t)^{S-1} (\bar{n}'_1 \bar{n}'_2 \bar{n}'_3 \cdots \bar{n}'_S)$.

6.5. The relationship between the description of a function in terms of its value at certain regularly spaced time intervals and in terms of its frequency spectrum as mentioned in the discussion of Sec. 6.3 is brought out by the Fourier-transformation theorem. Using the expressions for the Fourier transforms of Sec. 1.4 and assuming that $f(t) = \sqrt{2\pi}/(2\delta t)$ for $|t| < \delta t$ and $f(t) = 0$ for $|t| > \delta t$, show that $g(\omega) = [\sin(\omega \delta t)]/(\omega \delta t)$,

which approaches unity for very small δt . If $f(t)$ consists of n functions of the above type spaced a time τ apart show that $g(\omega) = \sin(n\omega\tau/2)/\sin(\omega\tau/2)$.

6.6. Consider a simple electric circuit at the temperature T which consists of a resistance R , which is not frequency-dependent, in series with a pure capacity C . Show by means of Eq. (6.14) that the square of the fluctuating potential appearing across C in the angular-frequency interval $\delta\omega$ is

$$\overline{\delta V_C^2} = \frac{2kTR}{\pi} \frac{\delta\omega}{1 + (\omega CR)^2}$$

Show that on integrating this over the whole range of ω the resulting value for $\overline{\delta V_C^2}$ is kT/C in agreement with the result that would have been obtained directly from Eq. (6.13).

6.7. From the definitions of the mean energy \bar{u} and the entropy S in terms of $\ln Z$, show that the entropy may be written

$$S = k \frac{\partial}{\partial T} (T \ln Z)$$

6.8. A circuit intended to measure the value of the electronic charge by measuring the root-mean-square value of the fluctuating shot-effect current consists of a diode connected in series with a battery and resistance R across which is a voltage amplifier having uniform amplification over a band of 10^4 cycles and no amplification outside this frequency range. When the thermionic current through the 10^6 -ohm resistor R is $50 \mu\text{a}$, it is found that the root-mean-square value of the fluctuating voltage across R is $40 \mu\text{v}$ in the 10,000-cycle band. Calculate the value of the electronic charge. Calculate the ratio of the resistor noise to the shot-voltage fluctuation assuming R is at room temperature (300°abs).

6.9. The suspension system of a permanent-magnet type of galvanometer consists of a very fine vertical quartz fiber which supports a light frame at its lower end consisting of two small horizontal permanent bar magnets with their centers on the fiber axis and a very small mirror for measuring the angular displacement of these magnets when a torque is applied to the system. Assume that the moment of inertia of the suspended system is entirely attributable to the magnets, each of which has a mass of 1 mg and a length of 0.5 cm . When the system is set into torsional oscillation the period is found to be 40 sec . Show that the restoring torque per unit displacement is $1.03 \times 10^{-13} \text{ newton-m rad}^{-1}$. A beam of light is reflected from the mirror attached to the suspended system and falls on a scale 3 m distant. If the system is at room temperature ($\approx 300^\circ \text{ abs}$), show that a root-mean-square fluctuation of the light spot of nearly 1.2 mm would be expected from thermal fluctuations alone.

6.10. Calculate the root-mean-square displacement to be observed along an axis for a small spherical particle of mastic suspended in water under the following assumptions: The particle is a sphere 10^{-3} mm in radius; the viscous force on a sphere moving through a fluid is $6\pi\eta av$, where η is the coefficient of viscosity, a is the radius, and v is the velocity; the temperature is 20°C ; and the time interval between observations is 100 sec . η for water is 10^{-2} poises (dyne-sec cm^{-2}) or 10^{-3} in practical units (newton-sec m^{-2}).

6.11. Following the discussion of Sec. 6.2, but assuming that occupancy of an energy interval Δu by a system is not the same throughout the energy scale but that in the neighborhood of the energy u_i there is a weighting factor g_i for the energy scale, as if there were a number g_i of equally likely physical situations corresponding to the u_i , show that P in Eq. (6.5) becomes $N! \prod_j (n_j!)^{-1} g_j^{n_j}$. Show further that Eq. (6.7)

becomes $n_i = g_i e^{\lambda} e^{\mu u_i}$, and that the rest of the discussion of that section is valid if the partition function Z is defined as $\sum_j g_j e^{\mu u_j}$.

6.12. By considering the rate of transfer of momentum between a confined aggregate of independent particles and the walls containing them, show that for particles of finite rest mass the pressure exerted on the walls is $\frac{2}{3}$ the kinetic-energy density of the aggregate of particles and that for photons the radiation pressure is $\frac{1}{3}$ the energy density within the enclosing walls.

6.13. Show by integrating the expression for the energy density of thermal radiation as a function of wavelength over the range from zero to infinity that the total energy density of thermal radiation is given by (*Stefan's law*)

$$U = \frac{8\pi^5 k^4}{15c^3 h^3} T^4 \quad \left(\text{given that } \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90} \right)$$

6.14. A volume V is divided into two volumes V_1 and V_2 by means of a partition. V_1 contains f_1 moles of a perfect gas and V_2 contains f_2 moles of the same gas. If the partition is removed, show that the change in Helmholtz free energy is

$$-\left[f_1 \ln \left(\frac{V_1 + V_2}{f_1 + f_2} \frac{f_1}{V_1} \right) + f_2 \ln \left(\frac{V_1 + V_2}{f_1 + f_2} \frac{f_2}{V_2} \right) \right] RT$$

Note that if $f_1/V_1 = f_2/V_2$ there is no change in the free energy.

6.15. A volume V is divided into two volumes V_1 and V_2 by a partition. V_1 contains gas A , and V_2 contains gas B at the same pressure. If the partition is removed, show that the entropy increases by

$$\frac{Rf_1}{V_1} \left[V_1 \ln \frac{V_1 + V_2}{V_1} + V_2 \ln \frac{V_1 + V_2}{V_2} \right]$$

where f_1 is the mole fraction originally occupying V_1 .

6.16. Consider a solution of Eq. (6.33) with a greater negative slope than -1.588 at the origin which intersects the x axis and represents a positive ion. Show by an application of the outer-boundary condition that the intersection with the X axis of the tangent to the curve at the point where it meets the x axis is $1 - n/Z$, where n is the number of electrons remaining in the ionic structure.

6.17. If a uniform magnetic induction \mathbf{B} is established in the neighborhood of an atom, the induced precession of the electron structure at the rate $\omega' = -e\mathbf{B}/2m$ gives rise to a magnetic induction \mathbf{B}' at the nucleus which reduces the net magnetic induction to which it is subject. This effect is known as diamagnetic shielding and has been calculated for a Thomas-Fermi atom by W. E. Lamb, Jr. [*Phys. Rev.*, **60**, 817 (1941)]. Assuming the spherically symmetric Thomas-Fermi atom, show that the circulating-electron current density induced by \mathbf{B} is $\rho v = -(e/2m)$, $\mathbf{B} \times \mathbf{r}$ and that this produces a magnetic induction \mathbf{B}' at the central nucleus given by

$$\mathbf{B}'(0) = -\frac{e\mathbf{B}}{3mc^2} \frac{1}{4\pi\epsilon_0} \int \frac{\rho'(r)}{r} dr = -\frac{e\mathbf{B}}{3mc^2} \varphi(0)$$

where $\varphi(0)$ is the electric potential at the nucleus due to the electrons alone. Using the series expansion given in the text, show that

$$\mathbf{B}'(0) = -3.184 \times 10^{-5} Z^{\frac{1}{3}} \mathbf{B}$$

GENERAL REFERENCES

- Coolidge, J. L.: "Mathematical Probability," Oxford University Press, London, 1925.
- Fowler, R. H., and E. A. Guggenheim: "Statistical Thermodynamics," Cambridge University Press, London, 1939.
- Fry, T. C.: "Probability and its Engineering Uses," D. Van Nostrand Company, Inc., New York, 1928.
- Gibbs, J. W.: "Elementary Principles of Statistical Mechanics," Longmans, Green & Co., Inc., New York, 1928.
- Gurney, R. W.: "Introduction to Statistical Mechanics," chaps. 1, 2, McGraw-Hill Book Company, Inc., New York, 1949.
- Kennard, E. H.: "Kinetic Theory of Gases," McGraw-Hill Book Company, Inc., New York, 1938.
- Keynes, J. M.: "A Treatise on Probability," The Macmillan Company, New York, 1921.
- Lindsay, R. B.: "Introduction to Physical Statistics," John Wiley & Sons, Inc., New York, 1941.
- Schroedinger, E.: "Statistical Thermodynamics," Cambridge University Press, London, 1946.
- Tolman, R. C.: "The Principles of Statistical Mechanics," Oxford University Press, London, 1950.

CHAPTER 7

ELEMENTARY PROPERTIES OF MATTER

7.1. Classical Properties of Matter. The large-scale physical phenomena which characterize the properties of matter are so diverse in nature and so complex in detail that the first problem in their analysis is the recognition of those simplest features of greatest generality which the first attempt at a theory must describe. Approximate relationships between observable quantities, such as the perfect-gas law, do not accurately portray the characteristics of any real sample of matter. However, they are of great practical utility in the establishment of certain necessary basic concepts and in the definition of prototypes in terms of which a systematic and ordered description can be formulated. The merit of these first and simplest approximate relationships is determined principally by the extent to which they are capable of supporting successive logical refinements increasing the validity of the description both in generality and detail. Thus the virial formulation of the equation of state of a gas, which is an extension of the perfect-gas law, greatly improves the quantitative correspondence of the theory with the phenomena exhibited by real gases. Successive refinements of the theory are inevitably accompanied by increasing complexity, which is acceptable to the extent that it provides at least a commensurate increase in utility. Proceeding in this way classical physics developed initially as a cell-like structure in which the fields of mechanics, hydrodynamics, elasticity, heat, electricity, magnetism, and light tended to be independent of one another with certain types of phenomena and basic concepts peculiar to each. Such subdivisions still form convenient categories in the organization of the study of physics because large groups of phenomena may thus be described with a minimum of hypotheses and a maximum of simplicity. These idealized descriptions are definitely bounded in their utility and of necessity lead to contradictions and inconsistencies in areas where the classical fields overlap. The *rigid body* dealt with in mechanics is inadequate for the description of phenomena at points of contact where elastic properties are essentially involved, and the *ideal fluid* of hydrodynamics is not adequate for real problems of flow or vibration propagation in liquids or gases.

The major achievement in physics during the past century has been the unification of these separate fields of observation through a description of them in terms of atomic theory. The concept of heat as a form of energy and the electromagnetic theory of light were preliminary unifying steps of great practical importance which paved the way for the broader and more basic unifications of the theory of relativity and the quantum theory. However, there is not as yet a completely satisfactory formulation of a unifying theory that embraces all phenomena. Also, the complexity of the process of translating atomic properties into their detailed consequences for the properties of matter in bulk is such that these properties cannot thus be deduced with the precision of direct measurement. The highest possible precision is of course frequently necessary for practical engineering purposes. In consequence of the need for precise characterization of the properties of matter and the desire to achieve the insight and unification provided by atomic and quantum concepts, one proceeds from both directions toward a junction in which the various constants, parameters, and coefficients long in use for the description of properties of matter in bulk can be described with the best possible accuracy as consequences of atomic properties.

Following this approach it is convenient to retain many of the categories and concepts of classical physics recognizing that these are not always well defined or capable of precise description. Thus a fluid is a form of matter which offers little resistance to change in shape, though viscous fluids grade continuously through plastics and glasses to crystalline solids which have approximately the properties of ideal elastic solids. Again, gases are fluids which exert relatively little resistance to a change in volume, though of course at the critical point there is no distinction between a gas and a liquid. It is also a convenience to consider mechanical, thermal, electro- and magneto-static, and electromagnetic properties of these substances somewhat separately, though of course they are all interrelated through the energy concept and the properties of the atoms making up the material.

The properties of the tenuous aggregate of atoms or molecules known as a gas can be understood more completely and described more simply than can those denser states of aggregation known as liquids and solids. The equilibrium mechanical and thermal properties of a gas are described in terms of the energy content and the equation of state relating the pressure, density, and temperature. These formulations determine such experimental quantities as the specific heat, the compressibility, and the coefficient of thermal expansion. The conditions of chemical equilibrium are also expressible in terms of these quantities. The electric and magnetic properties of gases in thermal equilibrium are also most conveniently stated in terms of the energy content and are found in tables as

electric and magnetic susceptibilities or dielectric constants and magnetic permeabilities. Nonequilibrium phenomena are much more difficult to describe in the language of thermodynamics or statistical theory, but the general nature of mechanical and thermal properties such as diffusion, viscosity, and thermal conductivity coefficients and the velocity of propagation and absorption coefficient of acoustic waves can be given an approximate quantitative formulation. The rates of chemical reactions can also be handled in an approximately quantitative manner. Finally, nonequilibrium electromagnetic phenomena such as the propagation and absorption of electromagnetic waves can be expressed in terms of atomic and molecular properties.

The denser states of aggregation such as liquids and solids are characterized by much greater energies of interaction between the constituent atomic or molecular systems. To a very good approximation for a gas, the energy is the sum of the energies of the isolated atomic constituents; extensive parameters are simply proportional to the number of atoms present, and scalar constants of proportionality relate the pressure and volume, the dipole moment and the electric field, the magnetic moment and the magnetic induction, etc. The much greater energy of interaction between the atomic constituents of a solid interrelates mechanical, thermal, and electromagnetic properties, and the theory becomes more phenomenological and less directly interpretable in terms of atomic or molecular characteristics. In general a solid is not isotropic and the stress must be related to the strain by a matrix equation and is accompanied by the development of electric and magnetic moments which assume major proportions in electrostrictive and magnetostrictive materials. Changes in crystal structure and the ordering of individual atomic components are accompanied by mechanical, thermal, and electromagnetic changes. The crystal must be considered as an irreducible system, the properties of which may be described in terms of atomic concepts and characteristics but which is not in any simple sense merely the aggregate of the constituent atoms. In a quantum-mechanical sense a crystal is a large atomic system with characteristic energy levels differing from those of its constituent atoms and described by wave functions which are determined as much by the symmetry of the crystal lattice as by the properties of the atoms composing it.

7.2. Perfect Gases. Specific Heats. A perfect gas is one which obeys the equation of state $pV = RT$. The average internal energy \bar{u} is independent of the density or the mean interatomic distance. This is quite a good approximation for most real gases at such high temperatures and such low densities that the classical energy-distribution law applies. The internal energy, however, need not consist entirely of translational energy but may include rotational, vibrational, and electronic-excitation

contributions and may also depend upon the presence of gravitational, electric, and magnetic fields.

As was seen in Sec. 6.2 the energy distributions among the different types of motion are independent of one another on classical statistics. Rewriting Eqs. (6.9) in terms of the temperature, the number of molecules with the energy u_j is given by

$$n_j = \frac{Ne^{-u_j/kT}}{\Sigma e^{-u_i/kT}} = -NkT \frac{\partial}{\partial u_j} (\ln Z) \quad (7.1)$$

where $Z = \Sigma e^{-u_i/kT}$ and

$$\bar{u} = \frac{\partial}{\partial \mu} (\ln Z) = kT^2 \frac{\partial}{\partial T} (\ln Z) \quad (7.2)$$

It is more illuminating to consider the various contributions to \bar{u} independently, and the subscripts t , r , v , and e will be used to refer to translational, rotational, vibrational, and electronic motion, respectively. Thus $u = u_t + u_r + u_v + u_e$, and the partition function is the product of factors of the form $Z_n = \sum_j e^{-u_{nj}/kT}$. The atomic or molecular heat

C_a is immediately derivable from \bar{u} since by definition it is the specific heat per atom or molecule, or the rate of change of \bar{u} with temperature:

$$C_a = \left(\frac{\partial \bar{u}}{\partial T} \right)_v \quad (7.3)$$

Nothing more need be said about the mean translational energy \bar{u}_t , for that was seen in Sec. 6.5 to be $\frac{3}{2}kT$. This also follows immediately from the equipartition law with the continuous variables p_x , p_y , and p_z upon which \bar{u} depends as $u_t = \frac{1}{2}m(p_x^2 + p_y^2 + p_z^2)$. If each of these has $\frac{1}{2}kT$ associated with it, the total mean energy is $\frac{3}{2}kT$. The translational contribution to the molecular heat is thus $\frac{3}{2}k$.

In general the energy separations of electronic energy levels of atoms or molecules are large compared with kT and, in consequence of the exponential factor, the population of upper states is very small. Since kT is about $\frac{1}{40}$ ev at ordinary temperatures, the population of a level 5 ev above the ground state would be proportional to e^{-200} which is a negligible quantity. However, at very high temperatures, say $12,000^\circ\text{abs}$ as in stellar atmospheres, a level 5 ev above the ground state would be populated to an extent of $e^{-5} \cong 10^{-2}$ which might well be appreciable. Also, the ground state may be a component of a close doublet, and in this case both levels could be populated to a significant extent at readily attainable temperatures. For instance, the ground state of the molecule NO is the lower of a doublet of separation 0.0153 ev. The partition function Z , consider-

ing these two states only, is then

$$Z = 1 + e^{-u_e/kT}$$

Where the zero of energy is taken at the lower state, the mean energy is, from Eq. (7.2),

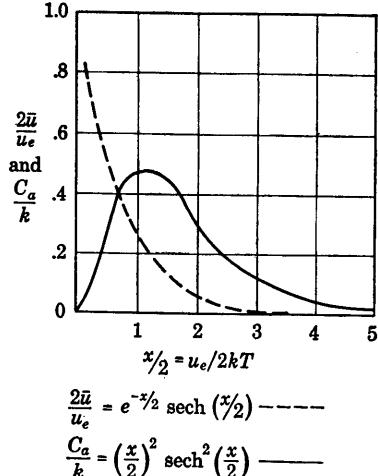


FIG. 7.1. Electronic-level contribution to the molecular heat.

and are important in specific-heat calculations in many gases, particularly the heavy ones. The population of the first excited vibrational level is proportional to $e^{-\hbar\omega_e/kT}$. Table 5.1 shows the exponential to vary widely for different molecules. Thus the contribution of this mode of motion

TABLE 7.1. POPULATION OF VIBRATIONAL STATES AT ROOM TEMPERATURE

Gas	$\tilde{\nu}_v, \text{cm}^{-1}$	$\hbar\omega_c/k$	$e^{-\hbar\omega_e/kT}$ for $T = 300^\circ\text{abs}$
H ₂	4395	6350	8×10^{-10}
O ₂	1580	2280	5×10^{-4}
Cl ₂	564	815	6.7×10^{-2}
I ₂	214	308	3.6×10^{-1}

to the specific heat would be expected to be quite appreciable for heavy molecules at room temperature. Proceeding to the calculation of the partition function for vibrational motion,

$$\begin{aligned} Z &= \sum_j e^{-(j+\frac{1}{2})\hbar\omega_e/kT} \\ &= e^{-x/2}(1 + e^{-x} + e^{-2x} + \dots) \\ &= e^{-x/2}(1 - e^{-x})^{-1} = \left(2 \sinh \frac{x}{2}\right)^{-1} \end{aligned}$$

¹ A. Eucken and L. d'Or, *Nachr. Gen. Wiss. Göttingen*, 107 (1932).

where $x = \hbar\omega_c/kT$. The level population is then given by Eq. (7.1) as

$$n_j = Ne^{-(j+\frac{1}{2})x} 2 \sinh \frac{x}{2} \quad (7.1')$$

and u_v by Eq. (7.2) as

$$\begin{aligned} u_v &= kT^2 \frac{dx}{dT} \frac{\partial}{\partial x} (\ln Z) \\ &= \frac{\hbar\omega_c}{2} \coth \frac{x}{2} \end{aligned} \quad (7.2')$$

and, finally, the molecular-heat contribution is

$$\begin{aligned} C_a &= \frac{\hbar\omega_c}{2} \frac{\partial(x/2)}{\partial T} \frac{\partial}{\partial(x/2)} \left(\coth \frac{x}{2} \right) \\ &= k \left(\frac{x}{2} \right)^2 \operatorname{csch}^2 \frac{x}{2} \end{aligned} \quad (7.4')$$

which resembles Eq. (7.4) except for the replacement of sech by csch. The curve of C_a as a function of x is shown in Fig. 7.2. It changes rapidly when x is about 3. Thus by reference to Table 7.1 it is evident that the vibrational levels cause a rapid rise in the specific-heat curve for hydrogen at about 2000° , for oxygen at about 700° , and for chlorine at about 250° abs.

Rotational levels, being in general more closely spaced, become important for specific heats at still lower temperatures. The general behavior of this contribution is a little different than that of vibrational motion because of the multiplicity of spatial orientations corresponding to a given rotational energy. If written in terms of the angular-momentum quantum number J , there are $2J + 1$ states with the energy $J(J + 1)\hbar^2/2M_rR_0^2$ (Sec. 5.3).

The partition function can be written in terms of the parameter J :

$$Z = \sum_J (2J + 1)e^{-J(J+1)\hbar^2/2M_rR_0^2kT}$$

This cannot be summed as conveniently as the previous expressions. However, for values of $M_rR_0^2$ and T such that $y = \hbar^2/2M_rR_0^2kT$ is small,

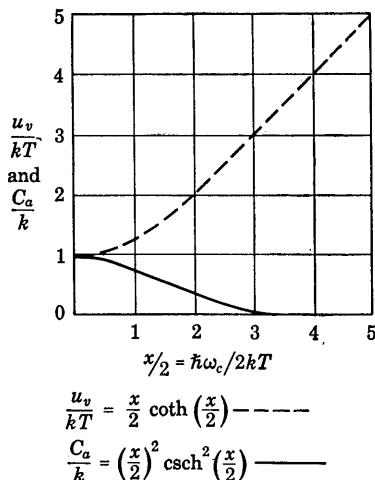


FIG. 7.2. Vibrational contribution to the molecular heat.

the sum may be approximated as an integral. Writing $J(J + 1)y = z^2$, $2z dz = y(2J + 1) \delta J$, and taking δJ as unity in the sum,

$$\begin{aligned} Z &\cong \frac{1}{y} \int_0^\infty 2ze^{-z^2} dz = \frac{1}{y} [-e^{-z^2}]_0^\infty = \frac{1}{y} \\ &= \frac{2M_r R_0^2 k T}{\hbar^2} \end{aligned}$$

From Eqs. (7.1) and (7.2)

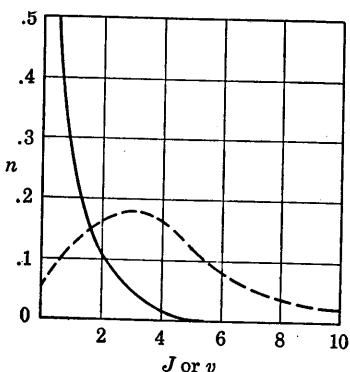
$$n_J = Ny(2J + 1)e^{-J(J+1)y} \quad (7.1'')$$

$$\bar{n}_r = kT^2 \frac{\partial}{\partial T} (\ln Z) = kT \quad (7.2'')$$

and the contribution to C_a is k . Equation (7.1'') is plotted in Fig. 7.3 for a representative value of y and is seen to differ from Eq. (7.1') in that there is a maximum at $2J + 1 = \sqrt{2/y} = (2R_0/\hbar)\sqrt{M_r k T}$. This is because of the weighting factor $2J + 1$. The population of the levels moves up with temperature in such a way that at sufficiently high

temperatures there may be a negligible number of molecules in low rotational states.

Equations (7.1'') and (7.2'') are, of course, not general, and the more interesting phenomena occur at temperatures for which y is not small. Writing y in terms of the wave number per centimeter to agree with Table 5.1, $\bar{\nu}_r = \hbar/2\pi c M_r R_0^2$,



Vibrational distribution

$$n_v \sim e^{-v}$$

Rotational distribution

$$n_J \sim y(2J+1) e^{-J(J+1)y} \quad \text{for } y = 0.04$$

For hydrogen, from Table 5.1,

$$\bar{\nu}_r = 121.62,$$

so

$$y_{H_2} = \frac{91.26}{}$$

FIG. 7.3. Rotational and vibrational distribution functions.

and y_{H_2} is not small for temperatures below several hundred degrees absolute. For O_2 , on the other hand, $y = 2.2/T$ and for all but the lowest temperatures the integral approximation is adequate and the rotational motion is fully developed. At temperatures below a few hundred degrees absolute one would expect ortho- and para-hydrogen to behave quite differently because odd rotational levels only are available to the former and even levels only to the latter (Sec. 5.5). Thus writing

the subscripts *o* and *p* for ortho- and para-, respectively

$$Z_o = 3e^{-2y} + 7e^{-12y} + 11e^{-30y} + \dots$$

$$Z_p = 1 + 5e^{-6y} + 9e^{-20y} + \dots$$

The molecular heats calculated from these expressions are plotted in Fig. 7.4 and are seen to be quite different below a few hundred degrees absolute. Since ortho-hydrogen has a spin of 1, there are $(2I + 1)$, or three, spatial orientations for the spin axis associated with each level, whereas for para-hydrogen $I = 0$ and the levels are single. Thus ortho-hydrogen has a statistical weight of 3 compared to para-hydrogen, and the partition function for the equilibrium concentration of the two types is $Z_{H_2} = Z_p + 3Z_o$. Putting the values of Z_o and Z_p in this expression, the internal energy and specific heat can be found for the equilibrium mixture. Also the equilibrium concentration Z_p/Z_{H_2} can be calculated as a function of the temperature. Maintenance of equilibrium depends of course on the presence of a catalyst that will permit the reaction $H_{ortho} \rightleftharpoons H_{para}$ to take place. The actual specific-heat curves for arbitrary concentrations are given by $a_o C_o + a_p C_p$, where a_o and a_p are numbers of ortho- and para-hydrogen molecules present in the mixture.

Finally, it should be remembered that the foregoing discussion relates to diatomic molecules only. Polyatomic molecules have an additional rotational degree of freedom and generally have a large number of different types of vibrational motion as well as partially free internal relative rotations between portions of the molecule. These lead to higher values of the specific heats for polyatomic molecules than for monatomic or diatomic molecules. Water vapor is an illustration of a gas whose molecule is not linear and which has three characteristic modes of vibration for which the wave-number separations are 5400, 5170, and 2290 cm^{-1} . At room temperature the three translational and rotational degrees of freedom each contribute $\frac{1}{2}kT$ leading to a molecular heat at constant volume of $3k$. The specific heat at higher temperatures reflects the increasing excitation of the internal vibrational degrees.

Magnetic and Electric Susceptibilities. The susceptibility of a substance is defined as the average induced moment per unit volume per unit inducing field. It is a very important practical parameter for a

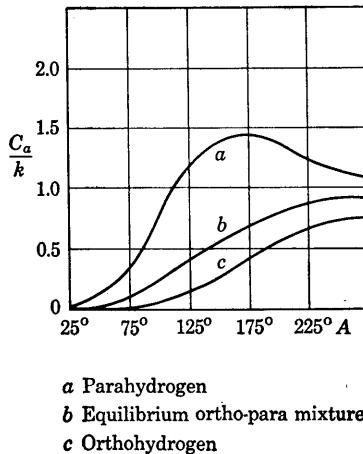


FIG. 7.4. Hydrogen molecular heat.

material, and it is closely related to atomic properties. In this section it will be considered only for those materials for which molecular interactions are negligible such as tenuous gases or sufficiently dilute solutions or suspensions. In conformity with the general definition of susceptibility, the magnetic susceptibility is defined as the induced magnetic moment per unit volume per unit magnetic induction \mathbf{B} . The magnetic susceptibility is intimately associated with the fundamental magnetic properties of individual atoms. Atoms frequently possess permanent magnetic moments (Secs. 2.6 and 4.1), whereas molecular combinations

TABLE 7.2. MOLECULAR HEAT OF WATER VAPOR, IN UNITS OF k

$T(A)$	$\omega_c(1)$	$\omega_c(2)$	$\omega_c(3)$	C_a	
				calc.	obs.
300	0	0	0.05	3.05	3.00
500	0	0	0.20	3.20	3.19
800	0.07	0.08	0.52	3.67	3.60
1000	0.15	0.16	0.65	3.96	3.90
1500	0.39	0.40	0.82	4.61	4.58
2000	0.57	0.59	0.90	5.02	5.04
3000	0.76	0.79	0.95	5.40	5.55

are usually of such a nature that the net moment is zero if the small nuclear contributions are neglected. The electric susceptibility is defined as the electric moment per unit volume per unit electric field \mathbf{E} . In distinction to the magnetic situation this quantity is not of particular fundamental significance though it is an important practical parameter. All atoms and many molecules are without any permanent electric moments, and the interest centers primarily on polar molecules.

When an atom is introduced into a region of magnetic induction \mathbf{B} , there are two primary effects. The first of these is that by Lenz's law the electronic motions are perturbed in such a way that a magnetic moment is induced which tends to counteract the inducing magnetic induction. The second is the tendency to orient any permanent magnetic dipole associated with the atom so that it has a net component parallel to the magnetic induction, as discussed in Sec. 2.6. The first effect is known as the diamagnetic effect and has its origin in the term neglected in the discussion of Sec. 2.6, and the second is known as the paramagnetic effect which, when present, is generally several orders of magnitude greater than the diamagnetic effect. The diamagnetic effect, which is always present and is independent of the temperature, may be thought of as arising from the induced precessional motion of the electronic structure at the rate $\omega' = (-e/2m_0)\mathbf{B}$ without regard to any permanent

magnetic moment associated with it. The magnetic moment of a system of charged particles is equal to the negative partial derivative of the Hamiltonian of the system with respect to the magnetic induction. Using Eq. (2.21) and assuming \mathbf{B} to be along the z axis, and further assuming that there are several electrons in each atomic system identified by the running index j ,

$$\begin{aligned}\mu &= -\frac{\partial H}{\partial B} = -\frac{\partial}{\partial B} \left[\sum_j \frac{1}{2m_0} p_j^2 - \frac{eB}{2m_0} (x_j p_{yj} - y_j p_{xj}) + \frac{e^2 B^2}{8m_0} (x_j^2 + y_j^2) \right] \\ &= \sum_j \left[\frac{e}{2m_0} (x_j p_{yj} - y_j p_{xj}) - \frac{e^2}{4m_0} (x_j^2 + y_j^2) B \right]\end{aligned}$$

The first term, which is the paramagnetic term yielding the component μ_p , has been discussed in Sec. 2.6. The second term is the diamagnetic one which may be put in terms of the mean square of the radius vectors to the charges of the system if a mean spherical symmetry is assumed such that $\bar{x}_j^2 = \bar{y}_j^2 = \bar{z}_j^2 = \bar{r}_j^2/3$. Thus

$$\mu_d = -\frac{e^2}{6m_0} \sum_j \bar{r}_j^2 B \quad (7.5)$$

This expression for μ_d can be evaluated say for the Thomas-Fermi atom and is found to be proportional to $Z^{1/3}$. The induced diamagnetic moment is independent of the position or orientation of an atom, and hence the diamagnetic susceptibility is unaffected by the random thermal motion.

The permanent atomic magnetic moments, if any are present, take up quantized orientations which have energies $-\mu_p \cdot \mathbf{B}$. This is a seat of energy which, in contrast to the diamagnetic energy, is in thermal equilibrium with the other types such as that of translational and rotational motion. The magnitudes of the moments and commonly available values of magnetic induction are such that this energy is small compared with the mechanical forms, but it is in general much larger than the diamagnetic energy. The average paramagnetic moment in the direction of \mathbf{B} can be calculated from the partition function if the magnetic energy is included in the latter, that is, $Z_m = \sum_j e^{-u_j/kT}$, where $u_j = -\mu_{p_j} B$ and μ_{p_j} is the component of the paramagnetic moment along \mathbf{B} for the energy u_j . Since

$$\begin{aligned}\mu_j &= -NkT \frac{\partial}{\partial u_j} (\ln Z_m) \quad \text{and} \quad \frac{\partial u_j}{\partial B} = -\mu_{p_j} \\ \bar{\mu}_p &= \frac{1}{N} \sum_j \mu_{p_j} n_j = kT \sum_j \frac{\partial u_j}{\partial B} \frac{\partial}{\partial u_j} (\ln Z_m) = kT \frac{\partial}{\partial B} (\ln Z_m) \quad (7.6)\end{aligned}$$

Following the discussion of Sec. 4.1 it will be assumed that the magneto-mechanical moment ratio, written γ , is $ge/2m_0$ and that the mechanical moment can have only the quantized components $m\hbar$ along \mathbf{B} , where m takes the values from J to $-J$. Thus, since $u_j = u(m) = \hbar\gamma m B$,

$$Z_m = \sum_{-J}^{+J} e^{-m\beta} \quad \text{where } \beta = \frac{\hbar B \gamma}{kT}$$

for any particular value of J . Assuming that all the atoms are in the rotational state having the angular momentum $\sqrt{J(J+1)} \hbar$, the partition function may be written as follows:

$$Z_m = \sum_{-J}^{+J} e^{-m\beta} = \frac{e^{(J+\frac{1}{2})\beta} - e^{-(J+\frac{1}{2})\beta}}{e^{\beta/2} - e^{-\beta/2}} = \frac{\sinh [(J + \frac{1}{2})\beta]}{\sinh (\beta/2)}$$

and from Eq. (7.6)

$$\bar{\mu}_p = \hbar\gamma \left\{ (J + \frac{1}{2}) \coth [(J + \frac{1}{2})\beta] - \frac{1}{2} \coth \frac{\beta}{2} \right\} \quad (7.7)$$

At the extreme of very low temperatures $\beta \rightarrow \infty$ and $\coth \beta \rightarrow 1$, so

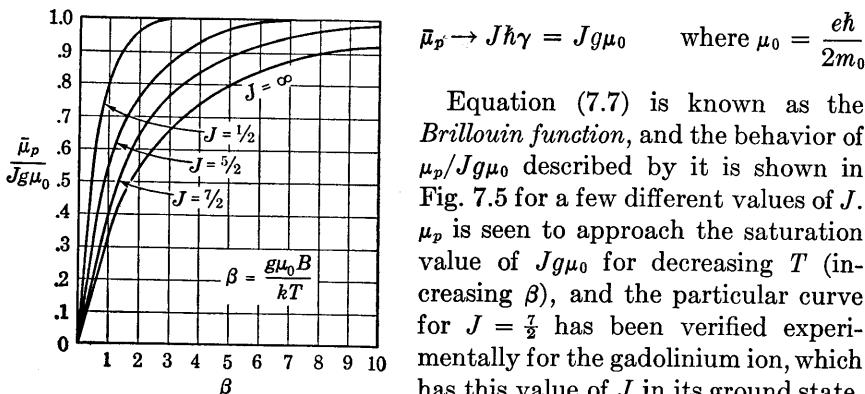


FIG. 7.5. Magnetic-susceptibility curves.

For ordinary values of T , β is very small, and to a good approximation for gases under common conditions the expansion $\coth x \rightarrow 1/x + x/3$ for small x may be used. Then Eq. (7.7) becomes

$$\bar{\mu}_p \rightarrow J(J+1)g^2 \frac{\mu_0^2}{3} \frac{B}{kT} \quad (7.7')$$

This is the *Curie law* for paramagnetism and represents the behavior of gases with permanent magnetic moments associated with their atoms.

¹ H. R. Woltjer and K. Onnes, *Communs. Phys. Lab. Univ. Leiden*, **167c** (1923).

$$\bar{\mu}_p \rightarrow J\hbar\gamma = Jg\mu_0 \quad \text{where } \mu_0 = \frac{e\hbar}{2m_0}$$

Equation (7.7) is known as the *Brillouin function*, and the behavior of $\mu_p/Jg\mu_0$ described by it is shown in Fig. 7.5 for a few different values of J . μ_p is seen to approach the saturation value of $Jg\mu_0$ for decreasing T (increasing β), and the particular curve for $J = \frac{7}{2}$ has been verified experimentally for the gadolinium ion, which has this value of J in its ground state, by Woltjer and Kamerlingh Onnes.¹

The susceptibilities due to the dia- and paramagnetic effects, χ_d and χ_p , are defined as $\mu N'/B$, where N' is the number of atoms per unit volume. The ratio of the relative magnitudes of dia- and paramagnetic effects for $J = 1$ is evidently

$$\frac{|\chi_d|}{|\chi_p|} = \frac{|\mu_d|}{|\mu_p|} = \frac{e^2 n a_0^2}{6 m_0} \frac{3}{2} \frac{kT}{g^2 \mu_0^2} \cong \frac{n}{g^2 \alpha^2} \frac{kT}{m_0 c^2}$$

where na_0^2 is written in order of magnitude for $\sum_j \bar{r}_j^2$ of Eq. (7.5). Since $m_0 c^2$ is of the order of 5×10^6 ev and α^2 is of the order of 5×10^{-5} , $\chi_d/\chi_p \cong nkT/25g^2$. For n and g of the order of unity and ordinary temperatures ($kT = \frac{1}{40}$ ev), the ratio of χ_d to χ_p is about 10^{-3} . Thus the paramagnetic effect, if present, generally completely masks the diamagnetic effect. An example of a molecule with a permanent magnetic moment is oxygen. The normal state of O_2 is a ${}^3\Sigma$ so that $g = 2$ and $J = 1$ and Eq. (7.7) leads to a molecular susceptibility of $8\mu_0^2/3kT$. This is in agreement with experiment for oxygen gas. The molecule NO is also paramagnetic, the normal state is ${}^2\Pi$, and here again the experiments¹ are in agreement with the theory.²

The effects associated with electric susceptibilities are of greater importance in fields of practical application such as the propagation of electromagnetic waves than are those associated with magnetic susceptibilities, but they are of less fundamental atomic interest. This is because the electrostatic energies involved are greater than the magnetostatic ones but the electric polarization is less intimately related to fundamental atomic properties than is the magnetic polarization. Indeed, the analogue of the paramagnetic effect, permanent electric moments, occurs only in polar molecules. The analogue of the diamagnetic effect, which is temperature independent, is the induced polarization as discussed in Sec. 5.1. It is of the opposite sign in the electric case, and the polarization resulting is in such a sense as to aid rather than oppose the applied field. The analogue of the temperature-dependent paramagnetic effect with the orientation of electric dipoles in an electric field is formally rather close, but the quantum-mechanical details are quite different. The motion of an electric dipole in an electric field is not necessarily a precessional one unless there is angular momentum associated with the molecule. However, in most cases the value of the mean induced dipole moment for gases assumes a form closely analogous to Eq. (7.7'). This can be derived semiclassically from the Boltzmann distribution law in

¹ E. C. Wiersung, W. J. De Haas, and W. H. Capel, *Communs. Phys. Lab. Univ. Leiden*, **212b** (1931).

² J. H. Van Vleck, *Phys. Rev.*, **31**, 587 (1928).

which $\mathbf{P}_0 \cdot \mathbf{E}/kT$ is inserted for u , where \mathbf{P}_0 is the permanent molecular dipole moment and \mathbf{E} is the electric field. The reason for this success is that the classical average value for the square of a vector is the same as the quantum-mechanical average. The vector in this case is \mathbf{P}_0 , and the mean value of the moment due to dipole orientation is $P_0^2 E / 3kT$ for high temperatures.

Thus the simple expressions for the magnetic and electric susceptibilities including both terms appear finally as follows:

$$\chi_m = \frac{N'}{B} \left(\mu_d + J(J+1)g^2 \frac{\mu_0^2}{3} \frac{B}{kT} \right)$$

$$\chi_e = \frac{N'}{E} \left(P_n + \frac{P_0^2}{3} \frac{E}{kT} \right)$$

Here P_n is the induced electric dipole moment. From these it is evident that the two terms can be determined separately by measuring the susceptibility as a function of the temperature. Then plotting χ as a function of $1/T$ yields the first term as the intercept of the line at $1/T = 0$ and the coefficient of $1/T$ in the second term as the slope of the line. The electric moments of the following molecules are representative:

TABLE 7.3. REPRESENTATIVE ELECTRIC DIPOLE MOMENTS OF POLAR MOLECULES

H_2O :	$P_0 = 1.85 \times 10^{-18}$ esu	$= 0.62 \times 10^{-29}$ coulomb-m
HCl :	$= 1.03 \times 10^{-18}$ esu	$= 0.34 \times 10^{-29}$ coulomb-m
HBr :	$= 0.78 \times 10^{-18}$ esu	$= 0.26 \times 10^{-29}$ coulomb-m
HI :	$= 0.38 \times 10^{-18}$ esu	$= 0.13 \times 10^{-29}$ coulomb-m

If the separated charge is taken as that of the electron, the separations are seen to vary from 0.21×10^{-10} m for HCl to 0.08×10^{-10} for HI. These are somewhat smaller than the Bohr radius a_0 as might be expected. Looking on the separation as related to the sum of the atomic radii, it is seen that the radius decreases from Cl to I qualitatively in accordance with the Thomas-Fermi theory of Sec. 6.5. It should be mentioned that more accurate values of the electric dipole moments of molecules have been obtained recently by Shulman and Townes¹ by a microwave absorption technique in the presence of an electric field. This method has a number of advantages, the principal one being that the moment for a particular rotational state is observed rather than the average over all rotational states. The difference between the absorption frequency in the presence and absence of the electric field is the difference in energy for the dipoles in the final and initial states in the field divided by Planck's constant.

¹ R. G. Shulman and C. H. Townes, *Phys. Rev.*, **77**, 500 (1950); R. G. Shulman, B. P. Dailey, and C. H. Townes, *Phys. Rev.*, **78**, 145 (1950).

7.3. Rate and Transfer Processes. It should be emphasized that the theory outlined in this and the preceding chapter relates only to the case of thermal equilibrium and does not apply to situations where the macroscopic phenomena are functions of the time. However, the success of the theory within this limitation and the great interest that attaches to the broader range of phenomena in which there is a change in the characteristics of matter with time or a net spatial transport of energy or momentum have led to attempts to make approximate extensions of the theory to include such phenomena. These attempts have been partially successful and yield very useful results, but they are fundamentally limited by the approximations that are involved and the lack of precision of certain of the concepts introduced. For instance, a concept which is basic in such an extension of the foregoing theory is the average number of collisions made by a molecule per unit time. But the concept of a collision between two molecules which have energies of interactions even at great distances is not capable of precise definition. If the molecules could be adequately approximated by spheres with sharp boundaries, the average number of collisions per second could be calculated, but such an approximation is not satisfactory and in consequence the value of such accurate calculations is seldom commensurate with the labor involved in making them. Here a simple and very approximate approach will be presented which yields results which agree with a more detailed analysis to within a factor of the order of unity.

Using the value of $e^{-\lambda}$ from Eq. (6.30) and the classical approximation to Eq. (6.18), the number of molecules per unit volume with momentum components in the range dp is

$$n'_p dp = \frac{4\pi n'}{(2\pi mkT)^{\frac{3}{2}}} p^2 e^{-p^2/2mkT} dp \quad (7.8)$$

Instead of using the convenient element of momentum-space volume $4\pi p^2 dp$ and integrating from 0 to ∞ , one could of course use $dp_x dp_y dp_z$ and integrate over each momentum coordinate from $-\infty$ to $+\infty$. By using this element of volume and performing two of the integrations, say y and z , since $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$

$$n'_{px} dp_x = \frac{n'}{(2\pi mkT)^{\frac{3}{2}}} e^{-p_x^2/2mkT} dp_x \quad (7.9)$$

This gives the number of molecules per unit volume with momentum components in the x direction in the range dp_x . These expressions can be written equally well in terms of the magnitude of the velocity p/m or the x component of the velocity as

$$n'_v dv = \frac{4n'}{\sqrt{\pi}} e^{-(v/v_m)^2} \frac{v^2 dv}{v_m^3} \quad (7.8')$$

$$n'_{vx} dv_x = \frac{n'}{\sqrt{\pi}} e^{-(v_x/v_m)^2} \frac{dv_x}{v_m} \quad (7.9')$$

where v_m is written for $\sqrt{2kT/m}$. The curves of n'_v/n' and n'_{vx}/n' are shown in Fig. 7.6. The maximum for v_x occurs at zero since positive

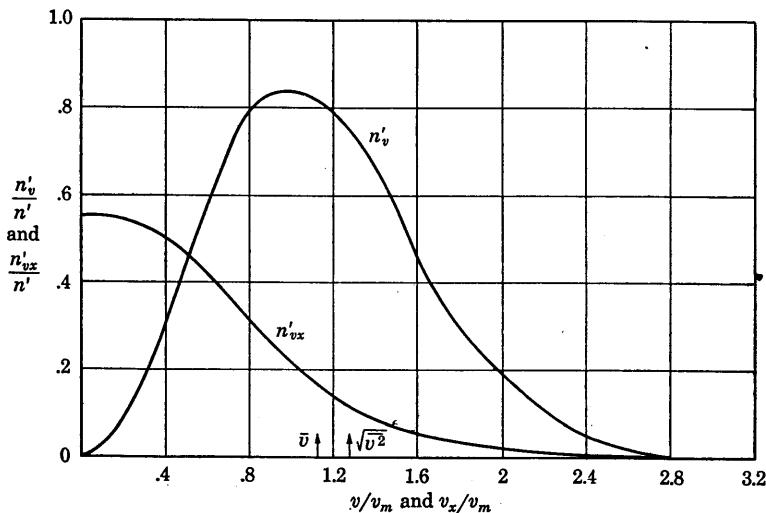


FIG. 7.6. Maxwell-Boltzmann velocity-distribution curves.

and negative values are equally likely. By differentiating Eq. (7.8'), it is seen that the maximum of n'_v occurs at v_m . At this point the velocity range dv contains the greatest number of molecules. The average velocity or mean-square velocity would be obtained by multiplying Eq. (7.8') by v or v^2 , respectively, integrating over the range of v , and dividing by n' . These values are indicated in Fig. 7.6 and are seen to be close to v_m but different from it and from each other. The differences between these quantities and the question as to which should be used in a particular type of calculation are representative of the uncertainties involved in the calculations here being considered.

The average distance traversed by a molecule between collisions, called the *mean free path*, can be readily estimated by considering a molecule

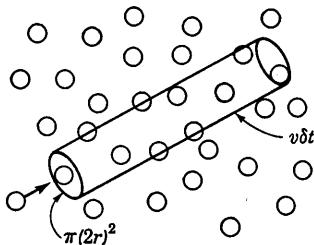


FIG. 7.7. Construction for discussing the mean free path.

moving among others which are at rest, as shown in Fig. 7.7. In a time δt it would move a distance $v \delta t$ and encounter any molecules whose centers lie within the cylinder of radius $2r$ and length $v \delta t$, where r is the radius of the molecules which are assumed to be spheres. Thus if n' is the number of molecules per unit volume, the moving molecule will make $\pi(2r)^2 v \delta t n'$ collisions in going the distance $v \delta t$, or the mean distance between collisions is given by $\lambda = 1/[\pi(2r)^2 n']$. Thus in terms of the molecular diameter $d = 2r$

$$\lambda = \frac{1}{n' \pi d^2} \quad (7.10)$$

A more careful analysis including the relative motions of all molecules, again considered as simple spheres, leads to a factor of $\sqrt{2}$ in the denominator. But the concept of mean free path is too indefinite for actual molecules to warrant much closer scrutiny.

Gas Reaction Rates. Some insight into the rate at which gas reactions would be expected to take place may be gained by a simple application of the preceding concepts. One obtains results in agreement with experiment if it is assumed that there is an activation energy u_a in the form of a potential-energy barrier between two approaching molecules that must be surmounted before they come close enough for the unit reaction to occur. The actual details of the energetics of the process are rather obscure, but the energy liberated at the reaction is transferred by radiation or direct encounter to the rest of the molecules of the gas. For the equilibrium concepts to be of any use as practical approximations, the rate of this energy redistribution among the molecules must be much greater than the rate at which the fast molecules, which are able by virtue of their high energy to react, actually do react with one another. Thus it would be expected that these considerations would be most applicable to slow reactions where only a very small fraction of the collisions result in reactions, and this is found to be the case.

As an illustration consider a reaction that takes place between molecules of type 1 and type 2 which occupy a common volume, the numbers of molecules of the two types per unit volume being n'_1 and n'_2 , respectively. The number of collisions per unit volume per unit time that occur for a molecule of type 1 for which the velocity component along a particular direction, in this case the line of centers, lies in the range of dv_x near v_x is $(v_x/\lambda) n'_{1v_x} dv_x$. The fraction of these collisions that occur with molecules of type 2 is $n'_2/(n'_1 + n'_2)$. So the number of collisions between molecules of these two types per unit time for which the relative velocity is sufficiently great to overcome the potential barrier to reaction

is by Eq. (7.9')

$$\begin{aligned} R_{12} &= \frac{n'_1 n'_2}{n'_1 + n'_2} \frac{v_m}{\lambda \sqrt{\pi}} \int_{v_{xc}}^{\infty} e^{-(v_x/v_m)^2} \frac{v_x}{v_m} \frac{dv_x}{v_m} \\ &= \frac{n'_1 n'_2}{(n'_1 + n'_2)} (n'_1 + n'_2) \sqrt{\pi} \frac{(r_1 + r_2)^2}{2} \left(\frac{2kT}{m}\right)^{\frac{1}{2}} \int_{\xi_c}^{\infty} e^{-\xi^2} d\xi \\ &= n'_1 n'_2 (r_1 + r_2)^2 \left(\frac{\pi k T}{2m}\right)^{\frac{1}{2}} e^{-u_a/kT} \end{aligned}$$

Here ξ is written for v_x/v_m and the activation energy u_a for $\frac{1}{2}mv_{xc}^2$, v_{xc} being the critical relative velocity and m the reduced mass. The simplest form of the mean free path is used, and the minimum of complication has been assumed. The result, however, is in satisfactory order-of-magnitude agreement with experiment, justifying the general correctness of the assumptions made. In the particular case that both the reacting molecules are of the same type, a reconsideration of the preceding argument shows that $n'_1 n'_2$ is to be replaced by $n'^2/2$, where n' is the molecular concentration. Thus, for instance, for the reaction $\text{HI} + \text{HI} \rightarrow \text{H}_2 + \text{I}_2$ studied by Kistiakowski¹ between about 500° and 800°abs , the coefficient of the concentration product $n'^2/2$ of R_{12} is found to be well represented by $3 \times 10^{-12} T^{\frac{1}{2}} e^{-22,000/T} \text{ cm}^3$ per molecule per second. Thus u_a is found to be 3.03×10^{-19} joule, or 1.9 ev, a very reasonable value. In practical units the first factor $(r_1 + r_2)^2 (\pi k / 2m)^{\frac{1}{2}} = 3 \times 10^{-18}$ or, using reduced mass for HI , $r_1 + r_2 = 0.5 \times 10^{-10} \text{ m}$, which is of the correct order of magnitude for atomic dimensions.

Viscosity. If a gas moves sufficiently slowly relative to some bounding surface, the gradation of the bulk velocity from an element of gas essentially in contact with the surface and at rest with respect to it to an element out in the gas stream so far away from the surface that the effect of the latter is negligible may be thought of as a continuous one. A representative molecule moving normally to the surface from one equilibrium region to the other must act as a transfer agent for momentum since its average momentum component must change by $u_x m$, where u_x is written for the velocity of the gas stream relative to the surface which is assumed normal to the z axis. The total rate of transfer of momentum by the molecules is of course the retarding force exerted between the gas and the surface. It is found under the simple conditions of a gas contained between two plane parallel surfaces of large extent, one of which moves parallel to the other, that the force of drag between the surfaces per unit area is proportional to the relative velocity divided by the separation. The constant of proportionality, written η , is known as the *coefficient of viscosity* of the gas. Consider the process as one of semiequilibrium in

¹ G. Kistiakowski, *J. Am. Chem. Soc.*, **50**, 2315 (1928).

which successive lamina parallel to the plates slip over one another with relative velocities $(du_x/dz) \delta z$, where δz is the lamina thickness. The net force on each lamina is zero in equilibrium, and the force per unit area on each plate, as illustrated in Fig. 7.8, is

$$F' = \eta \frac{du_x}{dz} \quad (7.11)$$

The coefficient of viscosity can be written in terms of molecular parameters by considering the rate of transfer of momentum from one plate to the other by the gas molecules and identifying the result with Eq. (7.11). Let $n'_{xz} dv_x dv_z$ be the number of molecules per unit volume with

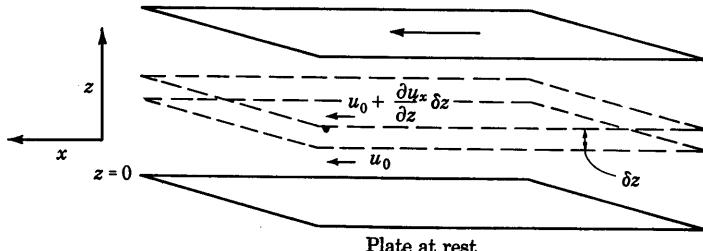


FIG. 7.8. Construction for viscosity discussion.

velocity components between v_x and $v_x + dv_x$ and between v_z and $v_z + dv_z$. Then $v_z n'_{xz} dv_x dv_z$ is the number of these crossing a unit area normal to z , and $m v_x v_z n'_{xz} dv_x dv_z$ is the rate at which they transfer the x component of momentum across a unit area normal to z . Thus

$$F' = \int_{v_z} \int_{v_x} m v_x v_z n'_{xz} dv_x dv_z = \eta \frac{du_x}{dz}$$

Now assume that the velocities in the x direction differ somewhat from the Maxwellian distribution of Eq. (7.9'); otherwise, by symmetry, F' would vanish. Thus let $v_x = v_{xm} + u_x$, where v_{xm} is the Maxwellian distribution [Eq. (7.9')], and $v_z = v_{zm}$, where v_{zm} is the Maxwellian distribution also. Then if $u_x = u_{x0} + (du_x/dz) \delta z$ near some fiducial lamina indicated by the subscript zero,

$$\begin{aligned} \eta \frac{du_x}{dz} &= \int_{-\infty}^{\infty} \int_{v_z}^{\infty} m v_{zm} (v_{xm} + u_0) n'_{xz} dv_x dv_z \\ &\quad + \int_{-\infty}^{\infty} \int_{v_z}^0 m v_{zm} \left(v_{xm} + u_0 + \frac{du_x}{dz} \delta z \right) n'_{xz} dv_x dv_z \end{aligned}$$

where the first integral is the x component of momentum transferred across the lamina from below and the second is that transferred across it

from above per unit time. Since v_{zm} and v_{xm} are independent and as often positive as negative, the contributions from such terms as $u_0 v_{zm}$ and $v_{zm} v_{xm}$ when integrated from $-\infty$ to $+\infty$ vanish and

$$\eta \frac{du_x}{dz} = m \frac{du_x}{dz} \delta z \int_{-v_x}^{\infty} \int_{-v_z}^{0} v_{zm} n'_{xz} dv_x dv_z .$$

The integral is the average value of v_{zm} for one sense only, and using Eq. (7.9') it is seen to be $n' v_m / 2\sqrt{\pi}$. So writing $\delta z = \lambda$, which assumes that the thickness of the lamina under consideration is of the order of the mean free path,

$$\begin{aligned} \eta &= \frac{mn'\lambda v_m}{2\sqrt{\pi}} = \frac{\rho\lambda v_m}{2\sqrt{\pi}} \\ &= \left(\frac{kTm}{2\pi^3 d^4} \right)^{\frac{1}{2}} \end{aligned} \quad (7.12)$$

where ρ is the gas density, and $v_m = (2kT/m)^{\frac{1}{2}}$ and $\lambda = (n'\pi d^2)^{-1}$ have been used. Equation (7.12) is quite successful in describing the behavior of gas viscosity in some ways but not in others. From Eq. (7.12), η should be independent of the pressure, and this is found to be the case. Also in accordance with Eq. (7.12), η is found to vary approximately as $T^{\frac{1}{2}}$. However, its dependence on m , the molecular mass, and d , the hard-sphere diameter, is not confirmed by experiment. This is not surprising as the hard-sphere picture is known to be a poor one. For further details the references dealing with kinetic theory should be consulted.

Heat Conduction. It is known that if a gas is contained between two extended parallel surfaces such as those of Fig. 7.8 and one of these surfaces is externally heated the other will rise in temperature at a greater rate than if the region between the plates were evacuated. This is due to the conduction of heat by the gas molecules or the net transfer of energy by the molecules from the hotter plate to the cooler one. As in the viscosity analysis the simplest assumption is that conditions are uniform over any lamina parallel to the plates, and the rate of transfer of energy can be considered as a function only of the gas, the plate separation, and the temperature difference. Then the net rate of energy transfer across any lamina can be written as

$$-w = K \frac{dT}{dz} \quad (7.13)$$

where the minus sign is used to indicate that the transfer is in the opposite sense to the temperature gradient and K is the *coefficient of heat conduction* of the gas.

Using the same concepts as in the viscosity discussion but considering the transfer of energy rather than momentum,

$$-w = \int_{v_z} n'_z v_z E_a dv_z = K \frac{dT}{dz}$$

where E_a is the energy carried by an atom or molecule. E_a may be expressed in terms of an effective molecular heat. Separating the translational and internal components,

$$E_a = C_a T \quad \text{where } C_a = \frac{mv^2}{2T_0} + C_{ai}$$

C_{ai} being the contribution from rotational or vibrational motion and T_0 a mean temperature. If the velocity distribution is approximately isotropic, the average of $v^2/3$ is approximately equal to the average of v_{zi}^2 . Assuming a Maxwellian distribution characteristic of the two sides of the lamina, the net rate of transfer of energy across it is

$$K \frac{dT}{dz} = \int_0^\infty n'_z v_{zm} \left(\frac{3mv_{zm}^2}{2T_0} + C_{ai} \right) T_0 dv_{zm} + \int_{-\infty}^0 n'_z v_{zm} \left(\frac{3mv_{zm}^2}{2T_0} + C_{ai} \right) \left(T_0 + \frac{dT}{dz} \delta z \right) dv_{zm}$$

where $T_0 + (dT/dz) \delta z$ with $\delta z = \lambda$ is the temperature to be used to a first approximation for molecules coming from above the fiducial lamina in a manner analogous to the treatment of viscosity. Just as in that treatment the terms cancel except for those due to $(dT/dz) \delta z$, and so

$$K = \lambda \int_{-\infty}^0 n'_z v_{zm} \left(\frac{3mv_{zm}^2}{2T_0} + C_{ai} \right) dv_{zm}$$

Using Eq. (7.9') these two terms may be integrated immediately, writing x for $(v_z/v_m)^2$, to yield

$$\begin{aligned} K &= \frac{\lambda n' v_m}{2 \sqrt{\pi}} \left(\frac{3mv_m^2}{2T_0} \int_{-\infty}^0 xe^{-x} dx + C_{ai} \int_{-\infty}^0 e^{-x} dx \right) \\ &= \frac{\lambda n' v_m}{2 \sqrt{\pi}} (3k + C_{ai}) \end{aligned} \quad (7.14)$$

This equation is closely related to Eq. (7.12), and on comparing them it is seen that

$$K = \frac{3k + C_{ai}}{m} \eta$$

This relation between K and η , which must be considered only approximate in view of the many simplifying assumptions, is, however, a very interesting one. The mean free path or molecular diameter is not

involved in it, and so rather good reliability can be anticipated. Table 7.4, of values taken from Chapman and Cowling¹ for a series of gases at 0°C, indicates the type of agreement found between the foregoing theory and observation. It is seen that there is an order-of-magnitude agreement between $Km/k\eta = KM/k\eta N$ and $3 + C_{ai}/k$ and that the trend with C_{ai} is correct.

TABLE 7.4. VISCOSITY AND HEAT CONDUCTION COEFFICIENTS FOR GASES

Gas	C_a/k (15°C)	K , joules m ⁻¹ deg ⁻¹ sec ⁻¹	η , newton-sec m ⁻²	M (mol. wt.)	$MK/\eta kN$	$3 + C_{ai}/k$
He	1.51	1.47×10^{-5}	1.887×10^{-9}	4	3.75	3
Ne	1.51	4.55×10^{-6}	2.981×10^{-9}	20	3.65	3
A	1.51	1.66×10^{-6}	2.104×10^{-9}	40	3.80	3
H ₂	2.46	1.74×10^{-5}	0.850×10^{-9}	2	4.9	4
N ₂	2.49	2.32×10^{-6}	1.674×10^{-9}	28	4.65	4
O ₂	2.54	2.45×10^{-6}	1.926×10^{-9}	32	4.85	4
CO ₂	3.42	1.47×10^{-6}	1.38×10^{-9}	44	5.65	5

7.4. Imperfect Gases. For a real or imperfect gas the state equation $pV = RT$ is not obeyed and the internal energy is not simply a function of the temperature. This implies that for real gases the free energy is not a linear function of the number of molecules. In accordance with the preceding discussion the free energies Ψ and Φ of N molecules were N times these functions for one molecule. Each molecule was thought of as occupying the entire volume independently of all the others. In a real gas there is an energy of interaction between each pair of molecules, and as a consequence there is a quadratic term in the expression for the dependence of energy upon number of molecules. This nonlinearity accounts for the greater portion of the complexity in any attempted description of an imperfect gas, vapor, or liquid. The present objective is to describe as simply as possible the effects that follow from an assumption of energy of interaction between molecules; for a rigorous discussion the references should be consulted.

The partition function for a single atom was defined in Sec. 6.2 as the sum of terms of the form $e^{-u_i/kT}$, where u_i represents the value of an energy state available to the atom and the a priori probability of occupancy of each energy state is assumed to be equal. As indicated in Prob. 6.11, if the a priori likelihood of occupancy of the energy state u_j is g_j , then the partition function is given by

$$Z = \sum_j g_j e^{-u_j/kT}$$

¹S. Chapman and T. G. Cowling, "The Mathematical Theory of Non-uniform Gases," Cambridge University Press, London, 1934.

If the system under consideration is not a single atom but consists of N independent atoms (as in a perfect gas), quantities such as the energies and entropies are additive. As a consequence of the proportionality between the energies and entropies and the logarithm of the distribution function, this function for the system is the product of the distribution functions for each atom of which the system is composed. Thus the *partition function* for all the atoms of the system is

$$\begin{aligned} Z^N &= \left[\sum_j g_j e^{-u_j/kT} \right]^N \\ &= \sum'_{u_j} \frac{N!}{\prod n_j!} \prod_j g_j^{n_j} \exp \left(- \sum_j \frac{n_j u_j}{kT} \right) \end{aligned}$$

where the primed sum is over all the product terms, each of which represents a particular distribution of the atoms among the energy states available to them consistent with the condition of constancy for the number of atoms: $\sum_j n_j = N$.

If there is an energy of mutual interaction between the atoms, the energy of a number of atoms is not simply the sum of the energies of them all separately, and hence the partition function for N atoms is not the product of the separate atomic partition functions. In general the energy states of an atom or molecule associated with translational, vibrational, and rotational motion or electronic excitation are independent of the presence of neighboring molecules, so the contributions of these modes to the partition function are of the simple product sort. On the other hand, there is in general a potential energy of relative position determined by the contiguity of molecular pairs. This is due to the field of force associated with each molecule and is responsible for the deviation of the equation of state from that of a perfect gas. The partition function determined by the energy levels arising from the mutual potential energy between molecules would by the definition of partition func-

tion also be written as $Z_p(N) = \sum_j g_{pj}(N) \exp \left(- \sum_j n_j u_{pj}/kT \right)$, but the u_{pj} 's are here functions of the n_j 's. The total partition function would then be the product of the Z 's for all types of energy states, or, writing Z_i for those factors that are independent of the number of molecules present,

$$Z(N) = (Z_i)^N \sum_{u_{ip}} g_{ip}(N) \exp \left(- \sum_j \frac{n_j u_{ip}}{kT} \right) = (Z_i)^N Z_p(N)$$

If the potential energy u_{jp} is simply a function of the separation from each molecule of all the rest and if uniform spatial distribution is assumed,

the summation in the exponent for each molecule is $-\sum_{j=1}^N u(r_j)/kT$, where

$u(r_j)$ is the energy of the j th molecule and the summation over u_{jp} is the multiple integral over all the pairs of molecules. If each integration is thought of in terms of the volume element $d\tau_j = 2\pi r_j^2 dr_j$, each $d\tau_j$ is weighted equally because of the uniform distribution, that is, $g_{jp}(N) = 1$ and

$$Z_p(N) = \int_{\tau_1} \cdots \int_{\tau_N} e^{-u(r_1)/kT} \cdots e^{-u(r_N)/kT} d\tau_1 \cdots d\tau_N$$

where the r 's with the subscripts are the separations of all other molecules from the one indicated by the subscript, and the $d\tau$'s are the elements of volume for each molecule. This is the product of factors of the form

$$Z_p(N)_j = \int_{\tau_j} e^{-u(r_j)/kT} d\tau_j$$

If $u(r_j) = 0$ the integration yields simply the volume V . If $u(r_j)$ is small the first two terms in the expansion will be an adequate approximation and $u(r_1)$, the energy of interaction of molecule 1 with all the rest, can be written as $(N - 1)\varphi(r)$, where $\varphi(r)$ is the interaction energy of a single pair separated by the distance r . Then

$$Z_p(N)_1 = V - (N - 1) \int_{r_1} \frac{\varphi(r_1)}{kT} 4\pi r_1^2 dr_1$$

The range of integration is from some small value representing the distance of closest approach to the largest value of r permitted in the volume. This may be taken as effectively infinite if $\varphi(r)$ becomes small sufficiently rapidly as r increases. Writing

$$\xi = \frac{-1}{V} \int_r \frac{\varphi(r)}{kT} 4\pi r^2 dr$$

$$Z_p(N)_1 = V[1 + (N - 1)\xi]$$

The second molecule interacts with but $(N - 2)$ others if none are to be counted twice, the third with $N - 3$, etc. So

$$Z_p(N) = V^N \prod_{n=1}^{N-1} [1 + (N - n)\xi] \quad (7.15)$$

If the second term in the parentheses is very small compared to unity, the product may be approximated as

$$1 + \xi \sum_{n=1}^N (N - n)$$

neglecting higher powers of the small quantity. The sum can be written as $\frac{1}{2}N(N - 1)\xi$, or approximately $\frac{1}{2}N^2\xi$, and using the further approximation that $\ln(1 + x) = x$ for small x

$$\begin{aligned}\ln Z_p(N) &= \ln V^N(1 + \frac{1}{2}N^2\xi) \\ &= N \ln V + \frac{1}{2}N^2\xi\end{aligned}\quad (7.16)$$

Since $\Psi = -kT \ln Z(N)$, $p = -(\partial\Psi/\partial V)$, and, as V appears only in $Z_{pot}(N)$,

$$\begin{aligned}p &= kT \left(\frac{N}{V} + \frac{1}{2}N^2 \frac{\partial\xi}{\partial V} \right) \\ &= \frac{NkT}{V} (1 - \frac{1}{2}N\xi)\end{aligned}\quad (7.17)$$

The second term in the parentheses is the correction to the perfect-gas law to take the forces between molecules into account.

If this equation of state is to be interpreted in terms of molecular properties, we must calculate ξ , which depends upon $\varphi(r)$, the energy of interaction of a pair of molecules a distance r apart. We may consider the problem in terms of the energy of interaction between two molecules polarized by one another's fields. The mutual potential energy of two similar molecules whose positive charges are a distance R apart and in which the negative charges have displacement coordinates x_1, y_1, z_1 and x_2, y_2, z_2 with respect to the positive charges as origins is

$$\begin{aligned}U_m &= \frac{e^2}{4\pi\epsilon_0} \left\{ \frac{1}{R} - [x_1^2 + y_1^2 + (R - z_1)^2]^{-\frac{1}{2}} \right. \\ &\quad + [(x_1 - x_2)^2 + (y_1 - y_2)^2 + (R - z_1 + z_2)^2]^{-\frac{1}{2}} \\ &\quad \left. - [x_2^2 + y_2^2 + (R + z_2)^2]^{-\frac{1}{2}} \right\}\end{aligned}$$

where e is the charge magnitude. The coordinates are indicated in Fig. 7.9.

If R is very large compared with the charge displacement, U_m may be expanded in powers of the displacement divided by R to yield approximately

$$U_m = \frac{e^2}{4\pi\epsilon_0 R^3} (x_1 x_2 + y_1 y_2 - 2z_1 z_2)$$

If there were no correlation between the displacements of the charges in the two molecules, this would of course average to zero, but as oscillations take place the displacements are correlated as would be the motions of two pendulums connected by a light spring.

The problem can be approached by the method of Sec. 5.4 or more directly by that of Sec. 5.1. The energy of a dipole of polarizability α

in a field E is

$$-W_e = \int_0^x eE dx = \int_0^E eE \frac{dx}{dE} dE = \int_0^E \alpha E dE = \frac{1}{2} \alpha E^2$$

since $ex = P = \alpha E$. Thus using Eq. (5.1) for the lowest-energy, or normal, state,

$$-W_e = \sum'_k \frac{|W'_{0k}|^2}{E_0 - E_k}$$

where the prime on the sum indicates that $k = 0$ is excluded. This is the value of W_e^1 for one dipole, and for a pair of atoms the energy is twice

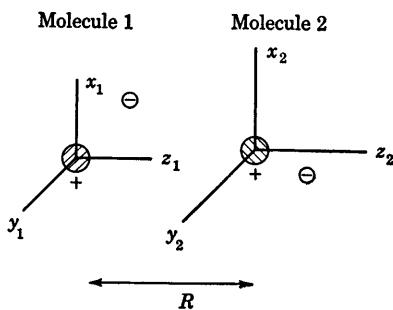


FIG. 7.9. Construction for van der Waals forces.

this. The proper calculation is difficult, and the references should be consulted for the details. However, a fairly satisfactory approximation results from a calculation of $\sum_k |W'_{0k}|^2$

divided by an order of magnitude value of $E_0 - E_k$. The latter will be taken as the ionization energy E_i of the valence electron. The sum may be simplified by making use of the orthogonality property of the wave function u_k . The function $U_m u_0$ can

be expanded in terms of u_k following the method of Sec. 4.6 or Sec. 5.1, and the coefficients of the u_k 's are seen to be W'_{0k} . This may be verified by multiplying the expansion by u_k and integrating, recalling that W'_{0k} is a constant:

$$U_m u_0 = \sum_k W'_{0k} u_k, \quad W'_{0k} = \int u_k^* U_m u_0 d\tau$$

So

$$\begin{aligned} \int u_0 U_m^2 u_0^* d\tau &= \int u_0 U_m (U_m u_0^*) d\tau \\ &= \int u_0 U_m \left(\sum_k W'_{0k} u_k^* \right) d\tau \\ &= \int \sum_k W'_{0k} u_0 U_m u_k^* d\tau \\ &= \sum_k |W'_{0k}|^2 \end{aligned}$$

This differs from the primed sum Σ' simply by the inclusion of the term $|W'_{00}|^2$ which is known by Sec. 5.1 to be zero. Thus, approximately,

$$\varphi(R) = 2W_e = \frac{-2e^4}{(4\pi\epsilon_0)^2 E_i R^6} \iint u_{01} u_{02} (x_1 x_2 + y_1 y_2 - 2z_1 z_2)^2 u_{01}^* u_{02}^* d\tau_1 d\tau_2$$

where the subscripts 1 and 2 refer to the two atoms. The cross-product terms vanish by symmetry, as with W'_{00} , and the quadratic term $x_1^2 x_2^2 + y_1^2 y_2^2 + 4z_1^2 z_2^2$ alone remains. The integral may be thought of as the average value of this quantity, and the average value of the square of each coordinate may be taken as one-third the average value of the square of the radius vector to the valence electron, r^2 , i.e.,

$$\varphi(R) = -\frac{4}{3} \frac{e^4}{E_i R^6} \frac{(\bar{r}^2)^2}{(4\pi\epsilon_0)^2}$$

Using this potential energy for a pair of atoms to evaluate ζ ,

$$\zeta = \frac{-1}{V} \int_R \frac{\varphi(R)}{kT} 4\pi R^2 dR = \frac{16\pi e^4 (\bar{r}^2)^2}{3E_i V k T (4\pi\epsilon_0)^2} \int_R R^{-4} dR$$

The upper limit of integration may be taken so great that its contribution $1/R_{\max}^3$ may be neglected, and the lower limit may be taken approximately as $(\bar{r}^2)^{\frac{1}{2}}$. Hence

$$\zeta = \frac{16\pi}{9E_i} \frac{e^4}{(4\pi\epsilon_0)^2} \frac{(\bar{r}^2)^{\frac{1}{2}}}{V} \frac{1}{kT}$$

or writing $\sqrt{\bar{r}^2} = Z^{-\frac{1}{3}} a_0$ from the Thomas-Fermi atom theory and $2E_0 = e^2/4\pi\epsilon_0 a_0$ from Sec. 3.4, where E_0 is the ionization potential of hydrogen,

$$\zeta = \frac{64\pi}{9} Z^{-\frac{1}{3}} \frac{E_0^2}{E_i k T} \frac{a_0^3}{V}$$

So to this approximation Eq. (7.17) becomes

$$p = \frac{NkT}{V} \left[1 - \left(\frac{32\pi}{9} Z^{-\frac{1}{3}} \frac{E_0^2 a_0^3}{E_i k T} \right) \frac{N}{V} \right] \quad (7.18)$$

$$= p_0 - A \left(\frac{N}{V} \right)^2 \quad (7.18')$$

where $A = (32\pi/9)Z^{-\frac{1}{3}}(E_0^2/E_i)a_0^3$ and $p_0 = NkT/V$.

Equation (7.18') is the equation of state, to the accuracy with which this analysis has been performed, for a gas whose molecules attract one another through the electric polarization that occurs between them.¹ It is by no means completely satisfactory, as many rough approximations have been made and the repulsive force between molecules has been entirely neglected. Still the results are in qualitative agreement with

¹ H. Margenau, *Rev. Mod. Phys.*, 11, 1 (1939).

experiment, and some insight is gained into the reason why the pressure of a real gas is less than that of a perfect gas as shown in Eq. (7.18'). The order of magnitude of the correction may be seen to be a few per cent at 0°C and 1 atm by putting in the values of the constants and taking E_i of the order of E_0 .

Though the above theory is intended only as a first approximation to the departure from the perfect-gas law, it is surprising how well it is in general agreement with observation if for V of Eq. (7.18') is substituted $V - Nb$, where Nb represents the volume actually occupied by the molecules and hence excluded from the region through which they may move freely. Evidently b is of the order of $\frac{4}{3}\pi a_0^3/Z$. Then from Eq. (7.18'),

$$(V - Nb) \left[p + A \left(\frac{N}{V} \right)^2 \right] = NkT \quad (7.18'')$$

to a first approximation if the correction terms to p and V are both small. Equation (7.18'') has long been known as van der Waals' equation, and it is reasonably successful in providing a general description of real gases.¹ The experimental observations are frequently given in terms of pV/NkT as a power series in N/V . Expanding Eq. (7.18''), neglecting the product of small terms ($N^3 b A / V^2$), and writing $p = NkT/V$ in the first small term,

$$\frac{pV}{NkT} = 1 + \left(b - \frac{A}{kT} \right) \frac{N}{V}$$

The units generally chosen in discussing this gas law are those in which the volume is measured in terms of the volume occupied by N molecules of the gas at 0°C and a pressure of 1 atm, i.e., units of $22.4 \times 10^{-3} \text{ m}^3$ (Amagat units). In terms of this volume and with p in atmospheres,

$$p_0 V_0 = NkT_0 = 1$$

and

$$pV = \frac{T}{T_0} + \frac{B}{V}$$

where B , known as the *second virial coefficient*, is $(bkT - A)N/22.4kT_0$. B has been measured for most permanent gases and is found to be approximately a linear function of T . The *Boyle temperature* is that for which B is zero, that is, A/bk . It is experimentally considerably below the value that would be predicted from values of b and A from atomic parameters, but it depends about as it should on $Z^{\frac{2}{3}}/E_i$ as may be seen from the table. The value of A is smaller than the calculated one indicating that the approximations were too crude, and the value of b

¹ J. A. Beattie and W. H. Stockmayer, *Repts. Progr. in Phys.*, 7, 195 (1940).

is larger than that calculated from molecular radii indicating that the hard-sphere picture is inadequate as was to be expected.

The foregoing outline of the theory of real gases indicates that their observed properties are in accord with predictions based on atomic properties and that the order of magnitude of the constants involved is correctly given by the atomic constants. The detailed comparison of the predictions of such an equation of state as Eq. (7.18') with measured Joule-Thomson coefficients, vapor pressures, etc., falls in the general field of thermodynamics and will not be taken up here. The qualitative

TABLE 7.5. IMPERFECT GAS PARAMETERS

Gas	Linear approximation for B	Boyle temp, deg. abs.
He	$-2.7 \times 10^{-5} + 1.42 \times 10^{-6}T$	19
Ne	$-3.75 \times 10^{-4} + 2.8 \times 10^{-6}T$	134
A	$-2.75 \times 10^{-3} + 6.8 \times 10^{-6}T$	410

agreement is much better than the crudity of the approximations might lead one to expect.

A further application of the theory of real gases is to the description of dilute solutions, which is well advanced from a phenomenological point of view. Dissociation, osmotic pressure, chemical reactions, etc., are all quite well accounted for on a picture of a comparatively inert matrix of solvent atoms or molecules in which are dispersed solute molecules which behave as if they formed a gas occupying the volume of the solution. The chief differences between a gas and a dilute solution are the reduction in electric forces caused by the dielectric constant of the solvent and the reduced rate of diffusion caused by the interference of solvent molecules with the free motion of solute molecules.

7.5. Approach to the Liquid State. As the pressure of a gas is increased or the temperature lowered, the average spacing of the gas atoms or molecules becomes smaller and smaller. At a characteristic temperature, called the boiling point, the density discontinuously increases and the molecules form a closely packed state in which each molecule is always interacting with several others. (If the temperature remains above a critical temperature, while the pressure is increased, this transition is accomplished without a discontinuous change in density.) The path of the molecules no longer contains regions in which a molecule is essentially independent of the others. Each molecule is simultaneously attracted to some others by the van der Waals, or polarization, force and repelled by other nearer molecules because of the exclusion principle which prevents overlapping of the orbital electron densities of the molecules (assuming they are chemically inert). Since the mean velocity of the

molecules depends only on the temperature, the energy change on condensation (latent heat of condensation) is a measure of the negative potential energy of interaction of the molecules. This binding energy is associated with the coalescence of the molecules into a drop with a free surface and a surface tension. As the temperature is lowered further, the liquid freezes into a solid. The distinction between a liquid and a solid is the presence of permanent order in the positions of the atoms in the solid. The transition between liquid and solid is usually associated with a discontinuous change in shear strength at the characteristic temperature of solidification. Nevertheless, glasses show only a gradual change in viscosity between liquid and solid phases. The atoms in a solid are constrained to a fixed position in the solid in contrast to the random movement of the atoms in the liquid. As a consequence, liquids are generally isotropic. Nevertheless, the liquid does exhibit some structure.¹ Scattering of X rays from liquids has shown a characteristic grouping of atoms which can best be described by a radial distribution function, i.e., the probability that a second atom occupy a volume element at a distance r from a selected atom. A peak in such a distribution function indicates a grouping called a *coordination shell*. The average number of atoms associated with this peak density is called the *coordination number*. Such a "most probable" distance implies a minimum in the interatomic potential-energy curve.

Measurements of the X-ray scattering from liquid argon will illustrate this concept.² Figure 7.10a shows the scattered X-ray intensity as a function of $s/4\pi = (\sin \phi/2)/\lambda$ from liquid argon at several pressures and temperatures. The analysis of these scattering measurements is accomplished much as in Sec. 5.1 with a determination of

$$\rho(r) = \rho_0 + \frac{1}{2\pi^2 r} \int_0^\infty s i(s) \sin rs ds$$

where $\rho(r)$ is the density of atoms at a distance r from a selected atom, ρ_0 is the average (over space) density of atoms, and $i(s)$ is the scattered X-ray intensity corrected for the effects of incoherent scattered radiation, atomic-structure factors, polarization of the incident beam, and absorption in the sample. The derived radial distribution curve for liquid argon is shown in Fig. 7.10b.

The prominent peak at 3.8 Å indicates the first coordination shell. The coordination number, or average number of atoms associated with this distance, varies from 10.5 at 84.4°K and 0.8 atm (near the solidification point) to 6 at 149.3°K and 43.8 atm (near the boiling point). A

¹ H. S. Green, "Molecular Theory of Fluids," p. 64, Interscience Publishers, Inc., New York, 1952.

² N. S. Gingrich, *Rev. Mod. Phys.*, **15**, 90 (1943).

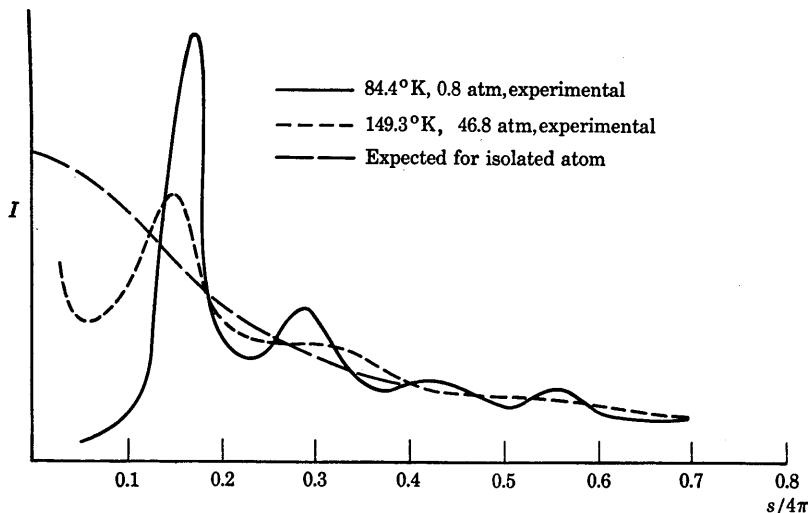


FIG. 7.10a. Intensity of scattered X rays from liquid argon. (Gingrich.)

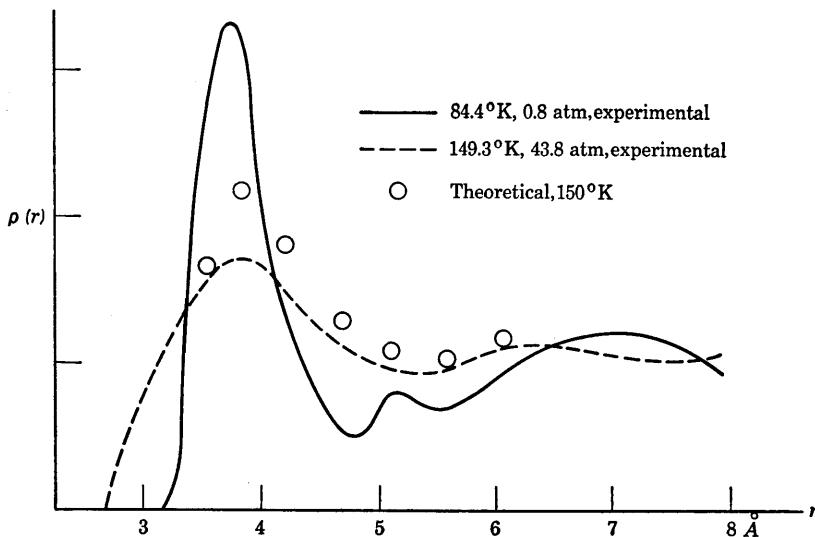


FIG. 7.10b. Radial distribution functions for liquid argon. (Green.)

second coordination shell appears at the lower temperatures, indicating about four atoms at a distance of about 5 Å. This is to be compared with the twelve nearest neighbors at 3.82 Å and six next-nearest atoms at 5.43 Å characteristic of solid argon which crystallizes in the face-centered cubic lattice. The gas phase retains a slight coordination of two atoms at 4.1 Å. These atomic regularities in the liquid phase imply no permanent order even of short range since the liquid is isotropic and the

atoms are continually moving about, changing position and producing density fluctuations.

The radial distribution function should in principle be calculable from a knowledge of the mutual potential-energy function of a pair of atoms. However, a consideration of the mutual influence of the neighboring atoms is quite complicated. An approximate calculation of this sort is discussed by Green,¹ and the results for argon at 150° are shown as circles in Fig. 7.10b. The agreement is encouraging, and, although the mathematical complexity inhibits a complete theory of liquids, various aspects are being clarified.¹

Scattering of Light. There are certain interesting phenomena associated with a material at its critical temperature and pressure. The *critical temperature* is that above which a gas cannot be liquefied by pressure alone, and the *critical pressure* is that at which the gas and liquid phases are in equilibrium at this critical temperature. An empirical equation of state such as Eq. (7.18'') is that of a surface, using p , V , and T as coordinates. The intersections of this surface with planes normal to T yield curves known as *isothermals*. Above the critical isothermal, which passes through the critical point, these curves are monotonic, but below it they are multivalued over a certain region of V which is associated with the liquid state. The critical isothermal is identified by having a point of inflection at the critical pressure and volume. The gas and liquid states are here specified by the same values of the variables, and the characteristic liquid surface is incipient.

It has been noted that near the critical point the mixture appears opalescent due to enhanced light scattering of a rapidly fluctuating or turbulent type. Though the critical parameters characterizing the point of inflection of Eq. (7.18'') are not in good agreement with the values predicted by the simple theory of A and b , the theory of thermal fluctuations previously developed does predict abnormal density fluctuations and hence abnormal light scattering at the critical point. This has been verified experimentally by Keesom,² who observed the light scattered at right angles from an unpolarized beam passing through ethylene near the critical point.

The intensity of light radiated by a dipole of maximum value P_0 oscillating with a frequency ω is given in Sec. 1.8. Taking the average value of the harmonic time factor as $\frac{1}{2}$ and the average of $\sin^2 \theta$ for the different polarizations at 90° observation as $\frac{1}{2}$, Poynting's vector is given by

$$\overline{N}_{\frac{\pi}{2}}(r) = \frac{\omega^4}{64\pi^2\epsilon_0 c^3} \frac{P_0^2}{r^2}$$

¹ H. S. Green, *loc. cit.*

² W. H. Keesom, *Ann. Physik*, **35**, 591 (1911).

The incident intensity in terms of the electric-field vector is $N_0 = \frac{1}{2}\epsilon_0 c E_0^2$. Thus the cross section for 90° scattering is

$$\sigma_{\frac{\pi}{2}} = r^2 \bar{N}_{\frac{\pi}{2}}(r) / N_0 = \frac{\omega^4}{32\pi^2 c^4} \left(\frac{P_0}{\epsilon_0 E_0} \right)^2 = \frac{\pi^2}{2\lambda^4} \left(\frac{P_0}{\epsilon_0 E_0} \right)^2$$

If a region of volume v_0 , as shown in Fig. 7.11, small compared with the cube of the wavelength, has a different number of atoms per unit volume in it than the uniform region surrounding it, it will scatter light because of the net polarization $P_0 = (\chi_e - \chi_{e0})v_0 E_0$, where χ_e is the volume electric susceptibility. As χ_e is proportional to the density,

$$P_0 = \frac{\rho - \rho_0}{\rho_0} \chi_{e0} v_0 E_0$$

and the scattering cross section of the volume is

$$\sigma_{\frac{\pi}{2}v} = \frac{\pi^2}{2\lambda^4} \left(\frac{\chi_{e0}}{\epsilon_0} \right)^2 \left(\frac{\rho - \rho_0}{\rho_0} \right)^2 v_0^2$$

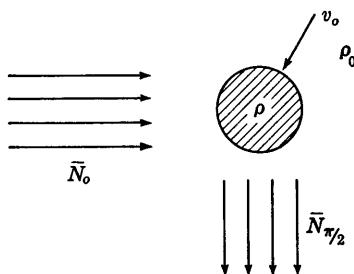


FIG. 7.11. Coherent light scattering from regions of varying density.

Finally, if the fluid is considered as made up of quite closely packed variations of density so that the number of these per unit volume is $1/v_0$, and they are in rapid random motion so that the light-scattering intensities add, the cross section for scattering per unit volume is

$$\sigma_{\frac{\pi}{2}v} = \frac{\pi^2}{2\lambda^4} \left(\frac{\chi_{e0}}{\epsilon_0} \right)^2 \delta^2 v_0 \quad (7.19)$$

where δ is written for $(\rho - \rho_0)/\rho_0$.

This equation shows in the first place that if $v_0 \ll \lambda^3$ the scattering of small λ 's will predominate. This is the cause of the blue of the sky and the red appearance of the sun on the horizon or through a haze. The average value of δ^2 can be calculated on the basis of thermal fluctuation theory. Consider a volume v_0 of an isothermal fluid that is compressed from the average surrounding pressure p_0 till it occupies the volume v_i . The work done, or the energy u_i stored in the volume, is then

$$u_i = - \int_{v_0}^{v_i} (p - p_0) dv$$

Assuming that p may be expanded in a Taylor series around the equilibrium pressure p_0 in terms of $(v - v_0)$,

$$\begin{aligned} p = p_0 + \left(\frac{\partial p}{\partial v} \right)_0 (v - v_0) + \frac{1}{2!} \left(\frac{\partial^2 p}{\partial v^2} \right)_0 (v - v_0)^2 \\ + \frac{1}{3!} \left(\frac{\partial^3 p}{\partial v^3} \right)_0 (v - v_0)^3 + \dots \end{aligned}$$

Integrating for u_j ,

$$u_j = - \left[\frac{v_0^2}{2} \left(\frac{\partial p}{\partial v} \right)_0 \delta_j^2 + \frac{v_0^3}{3!} \left(\frac{\partial^2 p}{\partial v^2} \right)_0 \delta_j^3 + \frac{v_0^4}{4!} \left(\frac{\partial^3 p}{\partial v^3} \right)_0 \delta_j^4 + \dots \right]$$

where $\delta_j = (v_j - v_0)/v_0$. Equation (6.12) then gives the number of volumes v_0 compressed to the volume v_j by thermal density fluctuations: $n_j \approx e^{-u_j/kT}$. The terms in u_j become successively smaller because of the assumed smallness of $v_j - v_0$, so if the first one exists it is the only one that need be considered. From it the average value of δ_j^2 , which is the same as δ^2 , is

$$\overline{\delta^2} = \frac{\int \delta^2 e^{\frac{v_0^2}{2kT} \left(\frac{\partial p}{\partial v} \right)_0 \delta^2} d\delta}{\int e^{\frac{v_0^2}{2kT} \left(\frac{\partial p}{\partial v} \right)_0 \delta^2} d\delta}$$

Integrating by parts this is immediately

$$\overline{\delta^2} = -kT \left[v_0^2 \left(\frac{\partial p}{\partial v} \right)_0 \right]^{-1}$$

which for a perfect gas, $p v_0 = n_0 k T$, is equal to $1/n_0$. Thus, for a perfect gas, Eq. (7.19) becomes $\frac{\pi^2}{2\lambda^4} \left(\frac{\chi_{e0}}{\epsilon_0} \right)^2 \rho_0^{-1}$. This expression yields quantitative agreement with scattered sunlight from the atmosphere.

The critical point is distinguished by the vanishing of $(\partial p/\partial v)_0$, and $(\partial^2 p/\partial v^2)_0$. Near this point, u_j becomes $\frac{v_0^4}{4!} \left(\frac{\partial^3 p}{\partial v^3} \right)_0 \delta_j^4$. The average value of $\overline{\delta_j^2}$ can be written as for the preceding case but with the different exponent yielding

$$\overline{\delta_j^2} = \overline{\delta^2} = \left[\frac{-(4!/v_0^4)kT}{(\partial^3 p/\partial v^3)_0} \right]^{\frac{1}{2}} \frac{\int_0^\infty x^2 e^{-x^4} dx}{\int_0^\infty e^{-x^4} dx}$$

The ratio of the integrals has the value 0.338, as may be found from tables of the gamma function. An assumption as to the equation of state determines the quantity in brackets. Taking van der Waals' equation (7.18'') and determining A and b by setting $(dp/dv)_c = (\partial^2 p/\partial v^2)_c = 0$ yields $(\partial^3 p/\partial v^3)_c = -\frac{27}{8} k T_c n_0 / v_c^4$ and $\overline{\delta^2} = 0.9 n_0^{-\frac{1}{2}}$. Thus the ratio of the light scattered near the critical point to that scattered by the gas far from this point is approximately $\sqrt{n_0}$, where n_0 is the number of atoms or molecules occupying a volume small compared with the cube of the wavelength of the light. In the case of ethylene, for instance, the critical density is 0.22 g cm^{-3} so there are about 47×10^6 molecules in a cube

1000 Å on a side. Thus the scattering at the critical point is about 10^3 to 10^4 times as great as at a higher temperature. It should be mentioned that the light scattered at the critical point is observed to be white rather than blue, indicating that the fluctuating volumes are no longer small compared with a wavelength of visible light and the simple theory here presented can be considered only approximate.

Liquid Helium. Another interesting instance of condensation in which the phenomena appear to be related to fundamental atomic properties is that of liquid helium. Helium is such a closely bound and symmetrical atomic structure that the interatomic forces may be considered negligible in comparison with those arising as a result of the statistical properties predicted for a high concentration of atoms. The experimental evidence is not as yet entirely conclusive, but it appears possible that a variety of liquid helium, known as helium II, which occurs below 2.17°abs and which exhibits very unusual properties may be accounted for by the Bose-Einstein type of statistics which the helium atom obeys since it has zero spin.¹

Equation (6.24) with $a = -1$ and $w = 1$ is the equation that determines the factor λ , and Eq. (6.25) then gives the mean energy for a Bose-Einstein distribution of helium atoms. At the opposite extreme to that of high temperatures and low densities for which the classical distribution law holds is the region of high densities and low temperatures associated with liquefaction or solidification. This corresponds to small negative values of λ , since if λ is positive the first term in the denominator can be less than unity for small values of x and a negative value of n' results which would be without physical meaning. Thus the limiting value of λ for high densities and low temperatures is zero. For such a small value of λ the assumptions upon which the sums for n' and \bar{u} were written as integrals no longer apply, but it is interesting to see what the consequences would be if Eqs. (6.24) and (6.25) were adequate. Taking $\lambda = 0$, the integrals involved in them become

$$\begin{aligned} \int_0^\infty \frac{x^2}{e^{x^2} - 1} dx &= \int x^2 e^{-x^2} (1 + e^{-x^2} + e^{-2x^2} + \dots) dx \\ &= \frac{\sqrt{\pi}}{4} (1 + 2^{-\frac{1}{2}} + 3^{-\frac{1}{2}} + 4^{-\frac{1}{2}} \dots) = 2.612 \frac{\sqrt{\pi}}{4} \\ \int_0^\infty \frac{x^4}{e^{x^2} - 1} dx &= \int x^4 e^{-x^2} (1 + e^{-x^2} + e^{-2x^2} + \dots) dx \\ &= \frac{3}{2} \frac{\sqrt{\pi}}{4} (1 + 2^{-\frac{1}{2}} + 3^{-\frac{1}{2}} + 4^{-\frac{1}{2}} + \dots) = \frac{3}{2} \frac{\sqrt{\pi}}{4} 1.341 \end{aligned}$$

¹ F. London, *Phys. Rev.*, **54**, 947 (1938).

where the sums are evaluated numerically. Therefore

$$\begin{aligned} n' &= 2.612 \left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} \\ \bar{u} &= 1.341 \frac{3}{2n'} kT \left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} = 0.513 \frac{3kT}{2} \end{aligned} \quad (7.20)$$

Thus the mean atomic energy of translation is about half of that for a gas under normal conditions. From the general relation that $pV = \frac{2}{3}\bar{u}$, and $n' = 1/V$ for a single atom,

$$p = 1.341 \left(\frac{2\pi m}{h^2} \right)^{\frac{3}{2}} (kT)^{\frac{5}{2}}$$

From this expression it is evident that p is independent of V and depends only on the temperature T in analogy with the vapor pressure of a liquid. Finally, the atomic heat at constant volume or constant n' is $(\partial\bar{u}/\partial T)_V = \frac{5}{2}\bar{u}/T = 1.92k$, about one-quarter greater than for a perfect gas.

Interest centers in the essential peculiarity of Eq. (7.20), which, since λ cannot be positive, states that the density n' of liquid helium cannot exceed the value there given. This is not an extremely great density at low temperatures, being only of the order of that of an ordinary liquid for $T = 4^\circ\text{abs}$. It would appear that pressures could be exerted which would increase the density well above such a value. This observation, together with the vapor-pressure-like dependence of p on T , suggests that some type of change of phase characterizes a Bose-Einstein gas for the small negative values of λ where the substitution of the integral for the sum in deriving Eq. (6.24) is not justified. Consider the sum of Eq. (6.18B), which in the integral form is Eq. (6.24).

$$N = \sum_j n_j = \sum_j \frac{g_j}{e^{-\lambda e^{u_j/kT}} - 1}$$

Though no n_j can be less than zero, the number of atoms occupying the lowest energy state, for which u_0 may be taken so small that the exponential term containing it approaches unity, can be made very large if $e^{-\lambda}$ is also close to 1. Taking λ as $-1/\alpha N$, where α is not small, the number occupying the lowest state is seen to be $g_0\alpha N$. In the case of radiation, which also obeys this statistics, g_0 vanishes because the weighting factor was seen in Sec. 6.4 to be $8\pi\nu^2 dv/c^3$ for radiation and this vanishes with ν . Such, however, need not necessarily be the case for an atomic aggregate. Thus, if g_0 and α are not too small, the lowest energy state could accommodate a considerable fraction of all the atoms and they may be thought of effectively as condensing into the analogue of a solid or liquid in that portion of the phase space described by the momen-

tum coordinates. The next-higher energy state would be much less densely populated as can be seen by considering that the order of magnitude of the energy difference δu between energy states must be that corresponding to the uncertainty principle for an atom located in a volume V . Thus as $\delta(mv) V^{\frac{1}{3}} \approx \hbar$ where $\delta(mv)$ is the momentum,

$$u_1 \approx \delta u \approx \frac{[\delta(mv)]^2}{2m} \approx \frac{\hbar^2}{2m} V^{-\frac{2}{3}} \approx \frac{\hbar^2}{2m} \left(\frac{n'}{N}\right)^{\frac{2}{3}}$$

If this is small, as it is known to be,

$$n_1 \approx \frac{g_1}{\left(1 + \frac{1}{\alpha N}\right) \left[1 + \frac{\hbar^2}{2m} \left(\frac{n'}{N}\right)^{\frac{2}{3}} (kT)^{-1}\right] - 1} \approx g_1 \frac{2mN^{\frac{2}{3}}}{\hbar^2 n'^{\frac{2}{3}}} kT$$

to the first order. And using n' from Eq. (7.20) as approximately adequate and α of the order of unity, $n_0/n_1 \approx N^{\frac{1}{3}}$ (see Fig. 7.12). This is very large for macroscopic samples, indicating the great population that may reasonably be expected in the lowest energy state in comparison with the higher ones. The integral formulation is adequate essentially for all but this very lowest state, and it may be that the strange physical properties of helium II are to be associated with the large occupancy of the lowest energy state at low temperatures and high densities. The extremely low kinetic energies associated with the lowest state lead essentially to nonlocalized atoms for collision purposes, which may account for the fact that the atoms of helium II in the lowest energy state appear unable to exchange momentum with those of helium I which may be thought of as composed of the atoms occupying the higher energy states.¹

7.6. Mechanical and Thermal Properties of Crystals. The physical phenomena exhibited by the solid state of matter are more tractable than those of liquids because of the high degree of permanent order exhibited by the constituent atoms or molecules. The principal distinguishing feature of a crystal is its translational-symmetry property. The atoms of a perfect or ideal crystal form a highly regular spatial lattice array such that three noncoplanar directions may be chosen along each of which the crystal may be displaced by any multiple of a characteristic elementary length and the same spatial distribution of all the atomic-

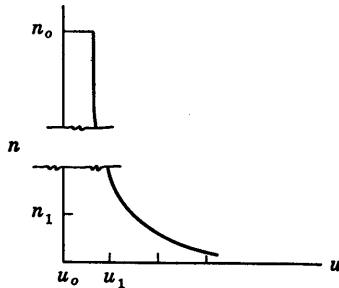


FIG. 7.12. Schematic of energy-level occupancy by a Bose-Einstein gas at low temperature.

¹ K. K. Darrow, *Rev. Mod. Phys.*, **12**, 257 (1940); K. Mendelssohn, *Repts. Progr. in Phys.*, **12**, 280 (1949); J. R. Pellam and P. R. Morse, *Phys. Rev.*, **78**, 474 (1950).

lattice intersections will be reproduced. The choice of these directions is not unique, though the axes are generally chosen in such a way that the elementary characteristic lengths are as small as possible. Crystals also may exhibit other types of symmetry than that of translation. The lattice pattern may be unchanged by rotations about certain axes or reflexion in certain planes; it is on the basis of such symmetries that the different crystal classes are defined. Real crystals differ from ideal ones by being finite in size, frequently distorted by strains, and having occasional vacant lattice sites, some foreign atoms, and some interstitial occupancy by atoms. The atoms do not occupy lattice points with mathematical precision but are held in place by elastic forces and vibrate about the lattice points as positions of equilibrium. There are also, generally, extensive surface imperfections, and the material may consist of small crystals imbedded in noncrystalline or amorphous material.

The interest of this discussion is primarily in the extent to which the characteristics of crystals can be accounted for by atomic properties. It is therefore most convenient, following Seitz, to consider crystals in categories based principally upon their electric properties. *Ionic crystals* such as NaCl are composed of electropositive and electronegative lattice centers, and in such crystals the small electric conductivity is predominantly due to the migration of positive and negative ions. These crystals show strong infrared absorption due to the vibratory motion of the ions occupying the lattice points. *Valence crystals* such as diamonds are held together primarily by homopolar or electron-sharing forces as opposed to the direct electric forces that predominate in ionic crystals. Valence crystals are generally very hard, their electric conductivity very low. *Molecular crystals* such as iodine, oxygen, rare gases, etc., are comparatively loosely bonded aggregates of atoms or molecules with little mechanical strength and low electric conductivities accompanied by small heats or energies of crystallization. *Semiconductors* such as silicon, germanium, or copper oxide are crystals which have electronic conductivities part way between those of insulators and metals and possess positive temperature coefficients of conductivity. Their conductivity is markedly affected by the presence of impurities or departures from stoichiometric proportions. Finally, there are the electropositive elements known as *metals* such as zinc or copper which have very high electronic-type conductivity and great mechanical strength.

The principal interatomic forces responsible for the regular ordering of a crystal lattice are electrostatic in origin. In general these are of the exchange type discussed in connection with molecules and are often known as valence forces. But in the cases of strongly electropositive and electronegative elements, which together form ionic crystals, the sharing of the electron is very one-sided, and to a good approximation

the valence electron of sodium may be thought of as actually attached to the chlorine atom in the NaCl lattice, for instance. Such cases are the simplest to consider in elementary terms, and, following Born, the forces for such crystals may be considered as derived from the electrostatic energy between lattice points having alternate charges $+e$ and $-e$ and close-range repulsive forces derived from the energy of the exchange potential of the form $Ae^{-r/\rho}$, where A and ρ are constants. Thus the energy represented by one ion pair is $Ae^{-r/\rho} - e^2/4\pi\epsilon_0 r$, where r is the separation of the pair.

The lattice of NaCl is of the simple cubic type as found by X-ray analysis, and the total energy of an ion is the sum of the energies for all pairs involving it, i.e.,

$$u(r) = \sum \left(Ae^{-r/\rho} \pm \frac{e^2}{4\pi\epsilon_0 r} \right)$$

where r takes the values for all possible choices of n_1 , n_2 , and n_3 in the expression $r = (n_1^2 + n_2^2 + n_3^2)^{1/2}l$, l being the distance of closest approach of the ions in the crystal. The negative sign is evidently appropriate when $n_1 + n_2 + n_3$ is odd, for then the ions are of unlike sign; the positive sign is appropriate when this index sum is even, for then the ions are of like sign. The short-range repulsive force may be neglected except for nearest neighbors, of which there are evidently six for each ion in a cubic lattice (see Fig. 7.13). The summation of the Coulomb term has been carried out numerically by Madelung¹ and shown to be equal to $-\alpha e^2/4\pi\epsilon_0 l$, where $\alpha = 1.7476$. Thus the energy of an ion in the cubic NaCl lattice is

$$u(l) = -\frac{\alpha e^2}{4\pi\epsilon_0 l} + 6Ae^{-l/\rho} \quad (7.21)$$

The above conclusion may be related to parameters such as the crystal compressibility. This is defined as $\chi = -\frac{1}{V} \frac{dV}{dp}$. The change in energy of one ion for a small change in l is, by the expansion of u near l_0 ,

$$\delta u(l) = \left(\frac{\partial u}{\partial l}\right)_0 \delta l + \frac{1}{2} \left(\frac{\partial^2 u}{\partial l^2}\right)_0 \delta l^2 + \dots$$

If the crystal is in equilibrium under its own forces, $(\partial u / \partial l)_0 = 0$. If then an external hydrostatic pressure δp is applied to the crystal, there

¹ E. Madelung, *Physik. Z.*, **11**, 898 (1910). See also M. Born, "Handbuch der Physik," 2d ed., vol. 24, part 2, p. 709. Julius Springer, Berlin, 1933.

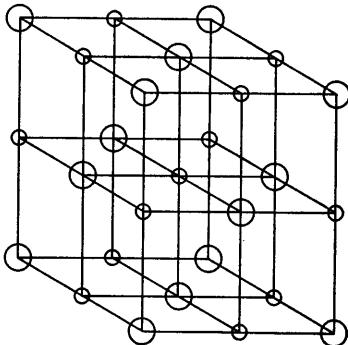


Fig. 7.13. Schematic of alkali-halide cubic lattice.

is a decrease in volume δV due to a shortening of all the atomic separations. If this change in volume is small, evidently $\delta V/V = 3 \delta l/l$ and the change in crystal energy δU is $\delta u(l)$ times the number of ions in the volume, or V/l_0^3 . Thus

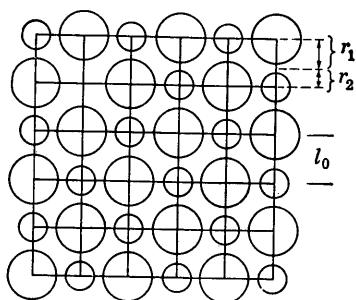


FIG. 7.14. Illustration of the parameters in the compressibility formula.

Finally, the condition of equilibrium $(\partial u/\partial l)_0 = 0$ can be used to eliminate A of Eq. (7.21):

$$\frac{\alpha e^2}{4\pi\epsilon_0 l_0^2} - \frac{6}{\rho} A e^{-l_0/\rho} = 0 \quad \text{or} \quad A = \frac{\rho \alpha e^2 e^{l_0/\rho}}{24\pi\epsilon_0 l_0^2}$$

and

$$u(l) = - \frac{\alpha e^2}{4\pi\epsilon_0} \left(\frac{1}{l} - \frac{\rho}{l_0^2} e^{-(l-l_0)/\rho} \right) \quad (7.23)$$

Also, the compressibility becomes

$$\chi = \frac{72\pi\epsilon_0 \rho l_0^4}{\alpha e^2 (2\rho - l_0)} \quad (7.24)$$

The agreement of this expression with experiment is quite good if allowance is made for the fact that l_0 , which is approximately the sum of the atomic radii, as in Fig. 7.14, is different for different atoms. The following empirical choice yields the best fit for the alkali-halide crystals:¹

TABLE 7.6. IONIC RADII FROM CRYSTALLINE ELASTIC PROPERTIES

Metal	$r_1, \text{ m} \times 10^{-10}$	Halogen	$r_2, \text{ m} \times 10^{-10}$
Li^+	0.475	Cl^-	1.475
K^+	1.185	Br^-	1.600
Cs^+	1.455	I^-	1.785

$$l_0 = r_1 + r_2, \quad \rho = 0.345 \times 10^{-10} \text{ m}$$

¹ N. F. Mott and R. W. Gurney, "Electronic Processes in Ionic Crystals," Oxford University Press, London, 1940.

It is also interesting to consider the ion energy [Eq. (7.23)] for an unstrained alkali halide ($l = l_0$):

$$u(l)_0 = - \frac{\alpha e^2}{4\pi\epsilon_0} \frac{l_0 - \rho}{l_0^2}$$

Taking the constants for KCl,

$$u_{\text{KCl}} = -1.32 \times 10^{-18} \text{ joules per ion}$$

In electron volts,

$$\frac{u_{\text{KCl}}}{e} = 8.25 \text{ ev}$$

This is the right order of magnitude for the formation of KCl as deduced from thermochemical data. In such a calculation the various energies such as the dissociation of Cl₂, the heat of crystallization of K, the ionization potential of K, and the electron affinity of Cl appropriate to the particular process taking place must be allowed for. The agreement for the same set of parameters for all the alkalis and halogens is not as good as can be obtained if less generality is attempted. Of course as only two parameters, ρ and l_0 , are concerned, perfect agreement can always be obtained from the compressibility χ and lattice energy $u(l)_0$ in the case of any one compound.

Specific Heat of a Crystal. The specific heat of a perfect gas is relatively easy to calculate as in Sec. 7.2 because each atom or molecule can be thought of as contributing independently to the energy content. This is clearly not the case for a crystal because of the large energy of interaction or coupling among all the constituent atoms. Except for imperfections the atoms may be considered as localized at the lattice points in elastic equilibrium under the influence of the crystal forces. These atoms can rotate or vibrate, or the electronic structure can be excited. Except in the case of relatively loose molecular crystal structures, rotational motion is not important, and, except for fortuitous circumstances such as a close doublet for the lowest electronic state, the electronic excitation term is not an important factor. The mean thermal vibrational energy makes the chief contribution, and this has been worked out for a harmonic oscillator in Sec. 7.2:

$$\bar{u} = \frac{\hbar\omega_c}{2} \coth \frac{\hbar\omega_c}{2kT}$$

The difficulty in applying this expression is in knowing what the characteristic frequencies are, and no completely satisfactory answer to this question has as yet been given.

In a crystal all the atoms at the lattice points interact with one another, and through these interactions the crystal is able to support a system of

elastic standing waves which resemble the system of electromagnetic radiation in an enclosure except that at the high-frequency end of the spectrum the pattern must be affected by the finite lattice spacing. Also, longitudinal as well as transverse waves are present. The simplest approximation is that which was proposed by Einstein in which the atomic interactions are neglected and each atom is considered as oscillating separately with three degrees of freedom. Then the total energy of vibration for N atoms in the crystal is

$$U = 3N\bar{u}$$

and

$$\begin{aligned} C_v &= \left(\frac{\partial U}{\partial T} \right)_v \\ &= 3Nk \left(\frac{\hbar\omega_c}{2kT} \right)^2 \operatorname{csch}^2 \frac{\hbar\omega_c}{2kT} \end{aligned} \quad (7.25E)$$

from Eq. (7.4'). The value to be used for ω_c comes from infrared experiments in which the resonant frequency is determined by the strong reflection from a crystal face of a particular frequency which is thus selected from a continuous background. The frequency that may be isolated in this way from black-body radiation by a series of successive reflections is known as the *reststrahlen* or the residual ray, and it is assumed to correspond to the frequency of the characteristic ionic motion in which the neighboring positive and negative ions vibrate out of phase under the influence of the electric vector in the light wave. Typical values for the wavelengths and angular frequencies for alkali-halide crystals are as follows:

TABLE 7.7. REPRESENTATIVE RESTSTRAHLEN FREQUENCIES

Crystal	Wavelength, m $\times 10^{-6}$	ω_c , sec $^{-1} \times 10^{18}$
NaCl	66.7	2.81
KCl	78.0	2.42
KBr	94.0	2.00
KI	115.0	1.64

An alternative and better approximation is that proposed by Debye in which the crystal is thought of as a continuum, like the region of space in the black-body-radiation analysis of Sec. 6.5, but the total number of modes of vibration possible to this region is limited by the lattice structure to the value $3N$ evidently imposed by the N particles. Thus there is a maximum frequency ν_m determined by

$$\int_0^{\nu_m} f(\nu) d\nu = 3N$$

where $f(\nu)$ is the number of normal modes of oscillation between ν and $\nu + d\nu$. Allowing for the fact that a solid can support longitudinal compressional waves in addition to the two polarizations of transverse shear waves and assuming that all these waves travel at the same velocity c , the appropriate weighting function $f(\nu)$ is $\frac{3}{2}$ times g_j of Sec. 6.5, or

$$f(\nu) d\nu = \frac{12\pi V}{c^3} \nu^2 d\nu$$

Then

$$U = \int_0^{\nu_m} f(\nu) \bar{u}(\nu) d\nu \quad \text{and} \quad 3N = \int_0^{\nu_m} f(\nu) d\nu$$

Neglecting the zero-point motion $h\nu/2$, as it makes no contribution to the variation of U with T , and since the mechanical vibrations, or *phonons*, obey the photon statistics,

$$U = \frac{12\pi V}{c^3} h \int_0^{\nu_m} \frac{\nu^3 d\nu}{e^{h\nu/kT} - 1}$$

and $3N = (4\pi V/c^3)\nu_m^3$. Therefore

$$U = \frac{9Nh}{\nu_m^3} \int_0^{\nu_m} \frac{\nu^3 d\nu}{e^{h\nu/kT} - 1}$$

The lattice specific heat at constant volume is then

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = 9Nkx_m^{-3} \int_0^{x_m} \frac{x^4 e^x dx}{(e^x - 1)^2} \quad (7.25D)$$

where $x_m = h\nu_m/kT$.

The dependence of (7.25E) and (7.25D) on T for very low temperatures is not the same. When $T \rightarrow 0$ the integral of (7.25D) approaches a constant $(\pi^2/90)^2$ and $\operatorname{csch} x \rightarrow e^{-x}$, so

$$C_{VE} \rightarrow 3Nk \left(\frac{h\nu}{kT} \right)^2 e^{-h\nu/kT} \quad (7.25'E)$$

$$C_{VD} \rightarrow 9Nk \left(\frac{\pi^2}{90} \right)^2 \left(\frac{kT}{h\nu_m} \right)^3 \quad (7.25'D)$$

for very small T . The experimental evidence strongly supports the cubic dependence of C_V on T predicted by the Debye form illustrated in Fig. 7.15. Also, for almost all crystals the behavior of the specific heat as a function of temperature follows very closely the form of Eq. (7.25D); the parameter for ν_m giving the best fit between theory and observation for typical metal crystals is given below:

Metal	Na	K	Cu	Ag	Au	Be	Zn	Hg
$h\nu_m/k$	150	100	315	215	170	1000	260	96

In the region of small T , where the crystalline vibrational specific heat is small, deviations are observed from the theoretical curves in a number of instances. The data for metals such as Al, Cu, and Ni, for instance, are best represented by a linear term in addition to the cubic term at very low temperatures, and this receives a reasonable explanation in terms of a free-electron contribution to the specific heat. In other cases such as Ge and Hf there are irregularities for small T which have not as yet been accounted for, and the reason for these departures may be due

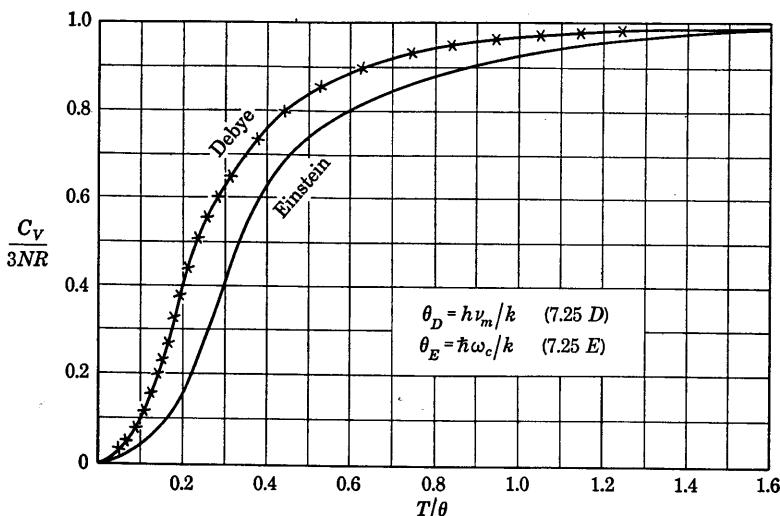


FIG. 7.15. Crystal-lattice specific-heat curves, x's indicate the nature of the experimental agreement.

to the inadequacy of the Debye approximation or some other unrelated cause.¹

7.7. Electron Band Theory of Crystals. Though it is not possible to formulate a completely quantitative theory of crystal properties on the basis of the characteristics of the constituent atoms, the principal features of crystalline behavior can be accounted for in a qualitative way and considerable insight gained by a simple consideration of the perturbations in the energy levels of an atom which are caused by the presence of its neighbors in the lattice array. It has been seen that the energy of an ion in a lattice is of the order of several electron volts, and this would be expected to have a major effect on the outer, or optical, energy levels of such an atom. The inner, or X-ray, levels would probably be but little affected by the presence of neighbors with interaction energies a hundred to a thousand times smaller than the energy of the X-ray levels. The outer levels, however, would be very considerably affected. Since the

¹ M. Blackman, *Repts. Progr. in Phys.*, **8**, 11 (1941).

coupling between all the atoms of the lattice is very close, the outer atomic levels may be thought of as being shifted variously in magnitude and constituting a very great number of levels characteristic of the crystal as a whole rather than of the individual atoms composing it. In the formation of a hydrogen molecule the ground states of the two hydrogen atoms become the attractive and repulsive states of the hydrogen molecule, the minimum of the former determining the nuclear equilibrium separation. In an exactly analogous way the ground states of the atoms composing a crystal perturb one another, leading to a number of crystal energy levels equal to the number of atomic levels times the number of atoms in the crystal. The energy minimum as a function of lattice spacing again determines the equilibrium configuration. A crystal 1 mm on a side would thus have a very large number ($\approx 10^{20}$) of energy levels occupied by the valence electrons of the constituent atoms. These levels differ from one another in energy due to the coupling between atoms just as the periods of coupled pendulums differ from the free periods. The energy range over which these levels occur is of the order of but a few electron volts, so the levels are very closely spaced and essentially form a band or continuum within the crystal. Thus a crystal such as that of sodium may be thought of as a lattice of singly charged sodium ions consisting of the nuclei, each surrounded by ten closely bound electrons, in $1s$, $2s$, and $2p$ states, and a continuum of energy levels pervading the crystal occupied by the $3s$ electrons from each of the individual atoms.

The problem is a three-dimensional one with the translational symmetry of the particular lattice determined by the energy minimum of the atoms concerned at the prevailing external conditions of temperature and pressure. It is not in general possible to calculate the particular structure assumed by the aggregate of atoms, but the commonest structures of simple atomic crystals are either those of closely packed spheres, i.e., hexagonal or face-centered cubic lattices, or the close variants of these. For an account of the approach to the three-dimensional problem, the special treatises on the subject should be consulted. Here only the simpler problem of an electron in a one-dimensional periodic perturbing field representing the equally spaced positive ions will be outlined.

The one-dimensional Schroedinger equation for an electron of effective mass m in a potential field V , from Sec. 3.4, would be

$$\frac{\partial^2 u}{\partial x^2} + \frac{2m}{\hbar^2} [E - V(x)]u = 0 \quad (7.26)$$

where $V(x)$ is the potential energy of the positively charged lattice ions arranged by regular spatial intervals, say l , along the axis of x . In the absence of $V(x)$ a solution would be a traveling wave $e^{\pm ikx}$. In the

presence of $V(x)$ it may be shown that a solution may be found of the form $u(x) = e^{\pm ikx}u'(x)$, where $u'(x)$ is periodic with the period l of the lattice, that is, $u'(x) = u'(x \pm l)$. As $u'(x)$ is periodic with the basic period l , it may be expanded in a Fourier series $u'(x) = \sum_{n=-\infty}^{\infty} A_n e^{-2\pi i n x/l}$. E and the coefficients A_n may be determined by substitution in Eq. (7.26) on the assumption that $V(x)$ is a small perturbing term. Choosing the positive sign for k and substituting the assumed u in Eq. (7.26),

$$\sum_{n=-\infty}^{\infty} A_n \left\{ - \left(k - \frac{2\pi n}{l} \right)^2 + \frac{2m}{\hbar^2} [E - V(x)] \right\} e^{-2\pi i n x/l} = 0$$

If $V(x)$ is small so that $u'(x)$ is not greatly dependent on x , A_0 , representing the free wave, will be the dominant term and $A_0 V(x)$ will be much larger than any other $A_n V(x)$. Therefore the equation may be written approximately as

$$A_0 \left\{ -k^2 + \frac{2m}{\hbar^2} [E - V(x)] \right\} + \sum_{n \neq 0} A_n \left[- \left(k - \frac{2\pi n}{l} \right)^2 + \frac{2m}{\hbar^2} E \right] e^{-2\pi i n x/l} = 0 \quad (7.26')$$

If this is multiplied by dx and integrated from 0 to l ,

$$A_0 l \left\{ -k^2 + \frac{2m}{\hbar^2} \left[E - \frac{1}{l} \int_0^l V(x) dx \right] \right\} = 0$$

since the periodic terms vanish. If the constant portion of $V(x)$ is assumed negligible, which involves no loss of generality, the integral must vanish since $V(x)$ is periodic with the spatial period l and the energy is

$$E = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m} \quad (7.27)$$

To this approximation the energy is thus the same as that of a free electron. Multiplying Eq. (7.26') by $e^{2\pi i n x/l} dx$ and integrating from 0 to l ,

$$-\frac{A_0 2m}{\hbar^2} \int_0^l V(x) e^{2\pi i n x/l} dx + l A_n \left[- \left(k - \frac{2\pi n}{l} \right)^2 + \frac{2mE}{\hbar^2} \right] = 0$$

or

$$A_n = \frac{2mV_n}{\hbar^2} \frac{A_0}{k^2 - (k - 2\pi n/l)^2} = A_0 \frac{l m V_n}{2\pi \hbar^2 n (k - \pi n/l)} \quad (7.28)$$

where $V_n = \frac{1}{l} \int_0^l V(x) e^{2\pi i n x/l} dx$, and the approximate value $E = \hbar^2 k^2 / 2m$ has been used.

Thus E and the coefficients A_n can be determined in terms of an arbitrary amplitude A_0 , but it is clear from the form of A_n that it is inconsistent to assume it small near values of the propagation vector \mathbf{k} which approach $\pi\mathbf{n}/l$, for here the denominator approaches zero. Thus the quadratic variation of E with inverse wavelength k is satisfactory except near values of $k = \pi n/l$ or $\lambda = 2\pi/k = 2l/n$. It is interesting to note that this corresponds to the one-dimensional Bragg condition of Sec. 2.2. In the general three-dimensional case the denominator of Eq. (7.28) is $k^2 - (\mathbf{k} - 2\pi\mathbf{n}/l)^2$ where \mathbf{k} and \mathbf{n} are vectors. This vanishes for

$$\mathbf{n} \cdot \mathbf{k} = \frac{\pi n^2}{l}.$$

As \mathbf{n} is normal to the Bragg planes and the diffraction angle $\phi/2$ of Sec. 2.2 is the complement of the angle between \mathbf{n} and \mathbf{k} , shown in Fig. 7.16,

$$\begin{aligned} nk \cos \theta &= nk \sin \frac{\phi}{2} = \pi \frac{n^2}{l} \\ \frac{2\pi}{\lambda} \sin \frac{\phi}{2} &= \pi \frac{n}{l} \\ n\lambda &= 2l \sin \frac{\phi}{2} \end{aligned}$$

which agrees precisely with the refraction condition for which propagation of the disturbance in a three-dimensional crystal is forbidden. The behavior of a disturbance near such a critical wavelength $k \cong \pi n/l$ can be looked at more closely by taking the particular A_n , which is large by virtue of this condition, specifically into account and neglecting all the others in the assumed solution of Eq. (7.26). Following this procedure, assume $u = A_0 + A_n e^{-ik'x}$, where $k' = 2\pi n/l$. This essentially selects the two terms considered large, A_0 and A_n , the latter by virtue of the near vanishing of the denominator of Eq. (7.28). Substituting

$$u = A_0 e^{ikx} + A_n e^{ik_n x}$$

in Eq. (7.26), where $k_n = k - k'$,

$$A_0 \left[-k^2 + \frac{2m}{\hbar^2} (E - V) \right] + A_n \left[-k_n^2 + \frac{2m}{\hbar^2} (E - V) \right] e^{-ik'x} = 0$$

Multiplying first by dx and integrating from 0 to l , and then by $e^{+ik'x} dx$ and integrating, the following two equations are obtained:

$$\begin{aligned} A_0(E - E_0) - A_n V_n^* &= 0 \\ -A_0 V_n + A_n(E - E_n) &= 0 \end{aligned}$$

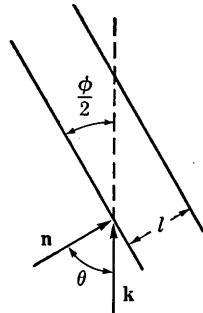


FIG. 7.16. Bragg reflexion parameters.

where $V_n = \frac{1}{l} \int_0^l V(x) e^{2\pi i n x/l} dx$, V_n^* is the complex conjugate of V_n , $E_0 = \hbar^2 k^2 / 2m$, and $E_n = \hbar^2 k_n^2 / 2m$. These homogeneous equations possess a solution if the determinant of the coefficients vanishes, that is, $(E - E_0)(E - E_n) - V_n V_n^* = 0$, or, solving this quadratic equation for E ,

$$E = \frac{1}{2} \{ (E_0 + E_n) \pm [(E_0 - E_n)^2 + 4V_n V_n^*]^{1/2} \}$$

In terms of k very near $n\pi/l$,

$$\begin{aligned} E &= \frac{1}{2} \left[\frac{\hbar^2}{2m} 2k^2 \pm 2(V_n V_n^*)^{1/2} \right] \\ &= \frac{\hbar^2}{2m} k^2 \pm |V_n| \end{aligned} \quad (7.29)$$

Thus the difference between the ascending curve from k_0 and the descending branch is the region of width $2|V_n|$ where there are no allowed energy values, as shown in Fig. 7.17. Putting this value in the equations for A_n and A_0 , $|A_n| = |A_0|$ and u , representing this forbidden region of E , is a standing wave and not one traveling through the crystal.

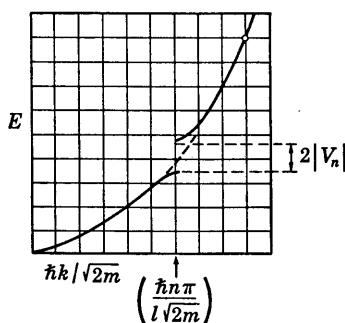


FIG. 7.17. Illustration of a one-dimensional forbidden band.

Electron Conduction. The proper three-dimensional case is of course considerably more involved than that of one dimension, but the general nature of the phenomena is similar. In the allowed energy regions known as *bands* the electrons act rather as if they were free, that is, $E = \hbar^2 k^2 / 2m^* = p^2 / 2m^*$. The electron energy is of course somewhat influenced

by the crystal lattice, and an asterisk is attached to the m to indicate that this may be thought of as a change in the effective mass brought about by the periodic lattice through which the electron moves. The density of energy levels, or the electron population of a band, depends of course on the basic nature of the atomic levels giving rise to it. One factor in determining the level density is that an *s* state contributes two levels per atom, a *p* state six, a *d* state ten, etc. To a first but frequently rather crude approximation the ion influences may be neglected and the density taken as proportional to the number of cells of the phase space corresponding to the spherical shell of radius p (momentum) and thickness dp , i.e.,

$$\frac{8\pi p^2 dp}{h^3} V = 4V\pi \left(\frac{2m^*}{h^2} \right)^{3/2} \sqrt{E} dE$$

In the region of the forbidden energies there are no levels to be occupied by electrons, but of course the Bragg condition depends upon the orientation of the vector \mathbf{k} with respect to the crystal axes. If the forbidden region $2|V_n|$ is large for all directions, it may be that in this neighborhood there is an absolutely forbidden region for all directions in the crystal and there is a finite gap between the permissible regions above and below. On the other hand, the gaps in all directions may not overlap and for one direction or another there may be permissible levels at all energies and electrons may be excited continuously through them to higher and higher levels without encountering an absolutely forbidden region.

This overlapping or nonoverlapping of the forbidden regions as in Fig. 7.18 is of particular importance in the electric conductivity of a crystal. For instance, in the case of a crystal composed of divalent atoms in normal s states, each double atomic state is filled with an electron for the isolated atoms. As the total number of levels remains unchanged when the atoms are brought together in a crystal, all the low levels are filled with electrons. If all higher levels, certain of which must be occupied if there is to be a net \mathbf{k} or drift in one direction, are separated by a gap in the level scheme, an electric field of moderate strength cannot raise electrons to such levels and no net flow of charge can take place. Such substances are then insulators. On the other hand, if the atoms are monovalent or if there is a continuum of energy levels above the electron population for divalent atoms, an external field can excite electrons into the occupancy of levels representing a net motion of charge and the crystal is an electric conductor.

A large number of different crystalline properties can be accounted for in a quantitative or semiquantitative way on the basis of the electron band theory. Approximate calculations can be made of the perturbations of the atomic energy levels as a function of lattice constant l . The way in which the upper energy levels of the copper atom, for instance, are affected by the increasing proximity of other copper atoms in a crystal is shown in Fig. 7.19. The broadening of the energy levels as l decreases is evident. When the atoms are very far apart the s , p , and d levels are very narrow as characteristic of a free atom. These levels then may be thought of as being broadened by the perturbation of the neigh-

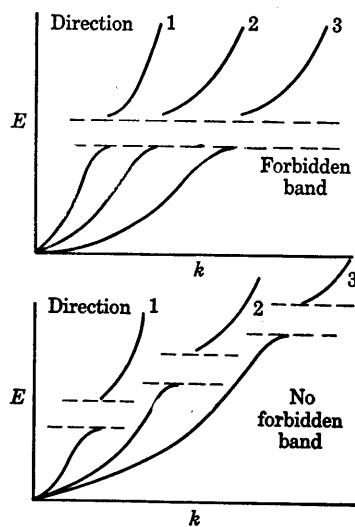


FIG. 7.18. Schematic of overlapping and nonoverlapping bands.

boring atoms as the separations are decreased until, at the actual equilibrium spacing l_0 , the 4s band covers a large energy span and the 3d band covers a smaller region within it. There are five times as many levels in the d band as in the s band (ten d and two s per atom), and hence the level density is much greater for d's than for s's. In the case of the copper atom the ground configuration is $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^1$. The single valence electron is in the s band and can move freely through the crystal in the conduction process. The ground configuration for the preceding atom, nickel, is $1s^2 2s^2 2p^6 3s^2 3p^6 3d^8 4s^2$, so the s band is full and one would

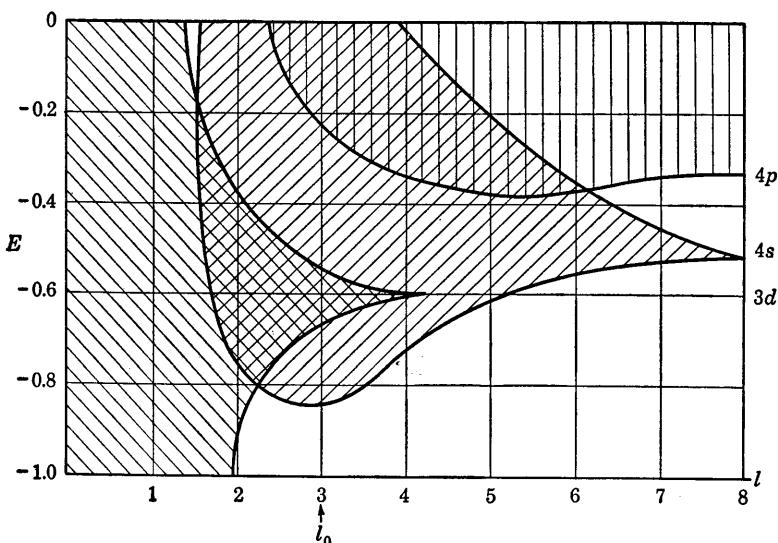


FIG. 7.19. Indication of the spread in energy of the electronic band structure in copper as a function of the lattice spacing.

expect conduction to be due to the freedom in the d band. Actually the perturbations of the energy levels alter this situation, and there is a certain fractional vacancy of electrons in both the s and d bands. The conductivity of nickel is less than that of copper presumably because of greater interference by the lattice ions with electron motion in the d band than in the s band.

Magnetism. The magnetic properties of metals are also closely related to the band structure. The diamagnetic effect need not be considered again because this contribution does not differ essentially from that described in Sec. 7.2. The paramagnetic effect is of course associated with the spin or orbital motion of the electrons; the closed shells contribute nothing because of the necessary symmetry in the moment orientations. The same may be said of filled nonoverlapping bands because, for instance, the two s electrons from each atom have necessarily oppositely oriented spins and hence no net moment. However, if there is but one s electron

(monovalency) or if the bands overlap, introduction of the specimen into a region of magnetic induction disturbs the equal occupancy of the lowest levels with exactly paired spins. This is illustrated schematically in Fig. 7.20. The band may be thought of as double, one half band consisting of spins with one orientation and the other half band of spins oppositely oriented. They are of equal energy extent in the absence of any external magnetic induction as shown by the dashed parabola, but in the presence of B one half band is lowered and the other is raised by the energy $B\hbar/2m_0$ of the electron moment in the magnetic induction B as indicated by the solid parabolas. If the bands are not full, electrons may transfer from one to the other by spin reorientation until the upper occupancy levels are of equal energy, i.e., the condition of equilibrium is attained. If $g(u)$ is the density of band levels near the top of the occupancy level of an energy scale, $g(u)B\hbar/2m_0$ electrons acquire a parallel position and the same number leave the antiparallel position. Since each of these contributes a moment $\hbar/2m_0$, the net moment per unit volume that develops is $2g(u)B(\hbar/2m_0)^2/V$, or the susceptibility χ_m (moment per unit magnetic induction per unit volume) is

$$\chi_m = \frac{2g(u)}{V} \left(\frac{e\hbar}{2m_0} \right)^2$$

The level density is much greater for the d than for the s band so elements with partially unoccupied d bands such as nickel would be expected to show much larger paramagnetic effects than those with filled d bands, such as copper, and this is found to be the case. To the crude approximation of neglecting lattice-ion influences $g(u)$ for a half band is, from the section on electron conduction, $2V\pi(2m^*/h^2)^{\frac{3}{2}}\sqrt{u}$. This may be expressed in terms of the number of electrons per unit volume in the crystal, and the energy value of the upper occupancy level by the condition $N/2 = \int_0^{u_m} g(u) du$ or $N/2 = 2\pi V(2m^*/h^2)^{\frac{3}{2}}\frac{2}{3}u_m^{\frac{3}{2}}$. Thus the level density near u_m is

$$\begin{aligned} g(u_m) &= 2\pi V \left(\frac{2m^*}{h^2} \right)^{\frac{3}{2}} \sqrt{u_m} \\ &= \frac{3}{4} \frac{N}{u_m} \end{aligned}$$

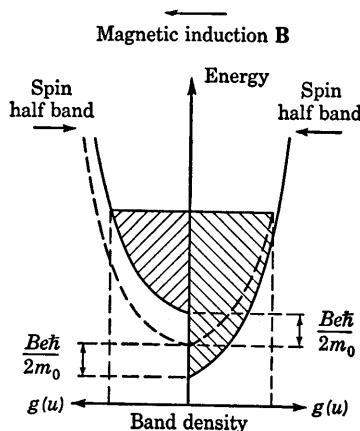


FIG. 7.20. Schematic indication of the development of paramagnetism by the orientation of electron spins.

and

$$\chi_m = \frac{3}{2} \frac{N'}{u_m} \left(\frac{e\hbar}{2m_0} \right)^2 = \frac{3}{2} \frac{N'}{u_m} \mu_0^2$$

where N' is the electron density per unit volume. This expression is in reasonable agreement with experiment.

A particularly striking feature of certain crystalline structures containing atoms or molecules having net electric or magnetic moments associated with them is the tendency for quantum-mechanical forces to induce a high degree of ordering in the orientation of these moments over macroscopic volumes, or domains. In these cases the forces of interaction between the electrons in the permeating band structure, as well as those more localized near the lattice points, and the ions composing the lattice are such that the minimum energy configuration is represented either by the lining up of the free moments associated with the lattice points along one of the crystal axes or by the alternate aligning of neighboring moments parallel and antiparallel to a crystal direction. The possibility of the latter configuration is of course contingent upon the lattice symmetry being such that the nearest neighbors of an atom are not also the nearest neighbors of one another. By hypothesis such highly ordered arrangements representing minimum internal energies for the crystal are assumed spontaneously in the absence of any external forces. The instance of uniform parallel alignment of magnetic moments is known as *ferromagnetism*. By analogy the parallel alignment of molecular electric moments in certain crystals is known as *ferroelectricity*. Crystals in which the forces are of such a nature as to induce the alternately parallel and antiparallel alignment of neighboring moments are known as *antiferromagnetic* and *antiferroelectric* for the two types of moments, respectively. The discussion here will be limited to the magnetic effects, which are the ones that have been studied most extensively. Excellent detailed accounts of the phenomena and the present state of the quantitative explanations of them are to be found in the references.¹

Ferromagnetic phenomena are observed only in the cases of crystals which contain atoms having partially filled internal electronic shells such as may be seen to be the case from Table 4.8 for the elements from scandium through nickel where 4s electrons are present but the 3d shell is incomplete. Other instances are the metallic groups further along in the table for which the 4d, 5d, and 4f shells are incomplete. The number of crystal structures containing such atoms that are ferromagnetic is comparatively very small because of the stringency of the condi-

¹ J. H. Van Vleck, *Rev. Mod. Phys.*, **17**, 27 (1945); C. Kittel, *Rev. Mod. Phys.*, **21**, 541 (1949); E. C. Stoner, *Repts. Progr. in Phys.*, **11**, 43 (1948); **13**, 83 (1950); R. M. Bozorth, "Ferromagnetism," D. Van Nostrand Company, Inc., New York, 1951.

tions of interaction between the electrons and the lattice ions necessary to bring about ferromagnetism. For the exhibition of ferromagnetic properties there must be labile net magnetic moments associated with the lattice points occupied by atoms with incomplete shells and there must also be energy minima associated with ordered alignments of these moments. The existence of electronic moments comparatively free to orient themselves under external forces and the identification of these with spin rather than orbital motion for atoms that display ferromagnetic phenomena are shown by two types of experiments. One of these is the direct measurement of the gyromagnetic ratio for a ferromagnetic crystal in which the application of a mechanical or magnetic torque causes the magnetomechanical moments to precess resulting in the establishment of a net angular momentum and accompanying magnetic moment along the axis of precession in accordance with the discussion of Sec. 2.6.¹ Another type of experiment is that in which the frequency of precession of these moments in the ambient magnetic induction is measured by determining the condition of resonance with an applied high-frequency electromagnetic field.² This is essentially the same technique as that described in Sec. 2.6 in connection with the measurement of nuclear moments. Such experiments establish the existence of these labile moments and also show that the g value is close to 2, thus identifying them primarily with the electron spin moment. The orbital motion is not entirely without influence but contributes only a few per cent to the total.

The identification of the moments with the lattice sites is most directly demonstrated by neutron diffraction experiments. It is found that neutrons are scattered by paramagnetic atoms both through the interaction between the neutron's magnetic moment and the magnetic moment of the extranuclear electronic system. When neutrons are scattered from a regular spatial lattice containing such atoms, the phase and amplitude of the wave representing the neutron are sensitive to the orientations of these moments. Neutron scattering from MnO was investigated by Shull and Smart³ at two temperatures. At the higher temperature, where any ordering of the moments would be expected to be greatly reduced by the thermal motion, diffraction due to the nuclear interaction was observed at angles characteristic of the chemical unit cell, with a continuous background due presumably to the random scattering. At the lower temperature, where the aligning forces should predominate,

¹ S. J. Barnett, *Proc. Am. Acad. Arts. Sci.*, **75**, 109 (1944); C. Kittel, *Phys. Rev.*, **76**, 743 (1949).

² W. A. Yager and R. M. Bozorth, *Phys. Rev.*, **72**, 80 (1947); W. H. Hewett, *Phys. Rev.*, **73**, 1118 (1948); W. A. Yager, *Phys. Rev.*, **73**, 1247 (1948); W. A. Yager and F. R. Merritt, *Phys. Rev.*, **75**, 318 (1949).

³ C. G. Shull and J. S. Smart, *Phys. Rev.*, **76**, 1256 (1949).

establishing ordered orientations of the moments over large domains, the diffraction maxima occurred at smaller angles characteristic of a unit cell twice as great as the chemical one and the continuous background was greatly reduced. These results are completely in accord with anti-parallel orientations of nearest manganese neighbors at the lower temperature, under which condition the recurrence interval of identical lattice points would be twice the spacing between manganese atoms. The anti-ferromagnetic property of MnO is also confirmed by measurements of susceptibility as a function of temperature. Scattering experiments such as these clearly identify the magnetic moments with the lattice points.

The most characteristic feature of ferromagnetism is the alignment of the spin moments associated with the lattice points over large crystal domains in the absence of any externally applied magnetic induction. The phenomena can be accounted for to quite a good approximation by adopting the hypothesis originally proposed by Weiss, in accordance with which a labile moment in a ferromagnetic material experiences an orienting force proportional to the average magnetic moment per unit volume in its neighborhood. Equation (7.7) gives the mean magnetic moment per atom in the presence of a magnetic induction B but in the absence of any forces of interaction between atoms. If to the orienting magnetic induction \mathbf{B} is added a field due to the effect of neighboring aligned moments, equal say to CI where I is the magnetic moment per unit volume, Eq. (7.7) would be written

$$\frac{I}{I_0} = \frac{J + \frac{1}{2}}{J} \coth [(J + \frac{1}{2})\beta'] - \frac{1}{2J} \coth \frac{\beta'}{2} \quad (7.7'')$$

where $\beta' = (g\mu_0/kT)(B + CI)$ and I_0 is the maximum value of I when all the moments are completely aligned. This equation with a constant value of J is not completely adequate to describe all ferromagnetic phenomena to the accuracy of present experimental results, but with $J = \frac{1}{2}$ as would be appropriate for a single electron with no orbital motion it furnishes a reasonable first approximation. Setting $J = \frac{1}{2}$ and $g = 2$, the equation becomes

$$\frac{I}{I_0} = \tanh \left[\frac{\mu_0(B + CI)}{kT} \right]$$

In the absence of any external magnetic induction, B vanishes and the equation is

$$\frac{I}{I_0} = \tanh \frac{I/I_0}{T/\theta} \quad (7.30)$$

where $\theta = \mu_0 CI_0/k$. Equation (7.30) has solutions only for $T < \theta$, but in this region it furnishes the basic description for the behavior of the

relative moment of a ferromagnetic crystal domain as a function of the temperature. It is plotted in Fig. 7.21 together with an experimental curve for a representative iron sample. θ is of the dimensions of temperature and is known as the *Curie temperature*. For temperatures less than θ the typical phenomena of ferromagnetism appear. For temperatures greater than θ a ferromagnetic substance is paramagnetic, and since for such temperatures the arguments of the coth functions are small, Eq. (7.7'') can be written

$$\frac{I}{I_0} = \frac{J+1}{3} \beta' \quad (7.31)$$

Using the values of g and J previously assumed and solving this equation for I/I_0 , it is found by comparison with the experimental variation of the atomic susceptibility that agreement is obtained with nickel, for instance, on the assumption that the magnitude of the moment per atom is $0.87\mu_0$. Below the Curie temperature the intensity of magnetization observed is in accord with a moment of $0.61\mu_0$ per atom. Discrepancies of this order are quite common in the comparison of magnetic parameters entering into the current theory and are doubtless to be attributed to the inadequacy of the description furnished by this theory of the very complex interactions between the partially mobile and partially constrained electrons and the lattice ion structures.

In the preceding discussion the orienting ferromagnetic force has been taken as proportional to the average magnetic moment per unit volume, but this does not imply that it is due to magnetic interaction between the electronic moments. Such magnetic forces are readily seen to be much too small to account for the Curie temperatures that are observed. The experimental values of θ for iron, cobalt, and nickel are 770°C , 1130°C , and 358°C , respectively, so $k\theta/\mu_0$ is of the order of 10^3 webers m^{-2} . This is too large by a factor of the order of 10^8 to be identified with CI if the force is to be of magnetic origin, since the largest observed magnetic inductions in permanent magnets are of the order of 1 weber m^{-2} . The energy corresponding to $k\theta$ is seen to be of the order of 0.1 ev and almost certainly arises from electrostatic interactions of an exchange type between the electrons, such as those discussed in connection with the helium atom and the hydrogen molecule. As illustrated in the case of helium the perturbing effect of one electron on another in a central field

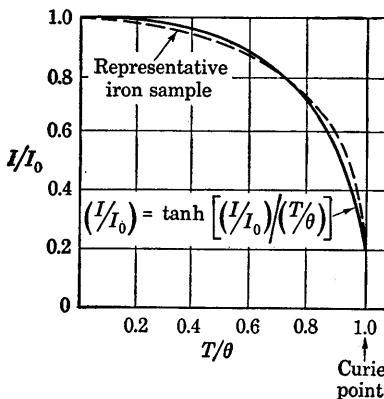


FIG. 7.21. Relative intensity of magnetization as a function of the ratio of the temperature to the Curie temperature.

favors separation of the electrons and parallelism of the spin moments. This is an instance of the general spectroscopic rule (Hund's rule) that within a multiplet the state of lowest energy is the one of highest multiplicity. Thus the tendency for the electrons in the unfilled d shell of an atom such as iron is toward a parallelism of their spins and a large resultant moment. On the other hand, as seen in the discussion of the hydrogen molecule, the perturbing effect of a second positive center on two electrons is to favor a closer approach of the electrons and, by the Pauli

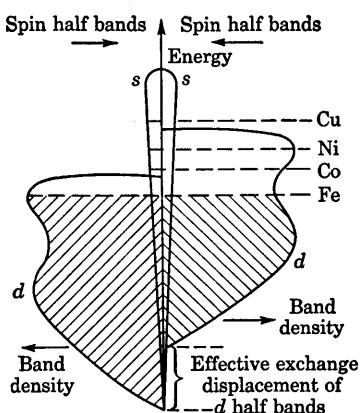


FIG. 7.22. Schematic indication of d -half-band displacements in ferromagnetism and occupancy levels for elements near nickel (iron cross-hatched).

lattice dimensions. However, it indicates how, as more electrons occupy the levels proceeding along the periodic table from iron to copper, the net magnetic moment per atom, which is proportional to the difference of the unoccupied levels of the two spin half bands, changes from 2.22 for iron to 1.71 for cobalt and 0.61 for nickel and vanishes for copper where the d half bands are both full.

The preceding discussion of the basic phenomena of ferromagnetism relates merely to the spontaneous establishment of a net magnetic moment along one of the crystal axes favored by symmetry over a small homogeneous domain. In general, a large polycrystalline sample does not exhibit a net moment because the moments of the mosaic of spontaneously magnetized domains which make it up are oriented at random. If an external magnetic induction is applied to the sample, those domains with magnetization components in the direction of this magnetic induction grow at the expense of those less favorably oriented. At first this is a comparatively continuous reversible process, and later, as larger crystal forces have to be overcome in order that the enlarged domains may shift

principle, a tendency toward oppositely oriented spins. It is not possible to carry through quantitative calculations of the results of such opposed exchange-force tendencies for the s and d electrons and the lattice ions of a crystal, but Fig. 7.22 is a highly schematic indication of the consequences of a relative displacement in energy of the two d spin-orientation half bands such as might result from electrostatic exchange forces. The figure is idealized in the simplicity of the energy level density contours of the s and d type bands, the assumption of uniformity of this structure throughout the crystal, and in its applicability to the four metals with differing atomic and crystal

their magnetization direction to the crystal axis most favored by the applied magnetic induction, the process becomes recognizably discontinuous and relatively irreversible in that the new directions of magnetization tend to remain after the removal of the applied magnetic induction. Finally, for very large magnetic inductions the directions of magnetization of the domains tend to rotate from the favored crystal axes of magnetization into parallelism with the direction of the applied magnetic induction. These three general stages in the development of magnetization in a crystal specimen are recognized in the initial toe, the steep rise, and the region of saturation of a typical magnetization curve.

X-ray Line Structure. The level density in bands has been approximated as proportional to \sqrt{u} in certain of the instances discussed, and it is interesting to note that a study of the shapes of X-ray absorption and emission lines gives direct experimental evidence on such level densities.¹ In the X-ray absorption process involving an atom in a crystal, an electron is raised by photon absorption from an inner atomic level, and the lowest-frequency photon that can effect this is one which raises the electron to the first unoccupied level in the electronic band structure. The probability of the process is proportional to the number of final states available to the electron, so a study of the shape of the low-frequency absorption edge yields information on the distribution in energy of levels above the normal electron occupancy level (see Fig. 7.23). Analogously, emission spectra result from the occupancy of a vacated X-ray level by one of the electrons already in the band above it. Again, since the probability of this process depends upon the number of electrons per unit energy range in the band, a study of emission-line shapes yields information on the level density within the occupied region of the band. In the soft X-ray emission bands of sodium it has been shown² that agreement with experimental results is achieved with the simple \sqrt{u} level density by taking into account the broadening of the energy levels in the conduction band due to electron collisions. Figure 7.24 indicates schematically the type of observed correspondence between level densities and X-ray-line shapes.

Thermionic and Photoelectric Effects. It is also interesting to see the extent to which a simple theory of a free-electron gas within the band,

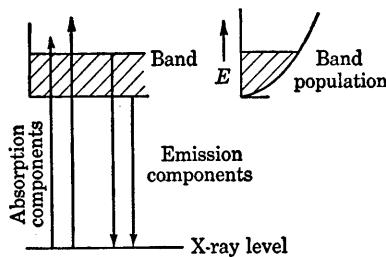


FIG. 7.23. Relation of X-ray-line shape to band levels and populations.

¹ H. W. B. Skinner, *Repts. Progr. in Phys.*, **5**, 257 (1938).

² P. T. Landsberg, *Proc. Phys. Soc. (London)*, **62A**, 806 (1949).

uninfluenced by any special lattice forces, can account for such phenomena as the external photoelectric effect and thermionic emission of electrons from conducting crystals.¹ In such quantum-statistical considerations one returns to Eqs. (6.24) and (6.25) to determine the mean energy u and other quantities of interest. Using appropriate parameters in Eq. (6.30) for a gas of free electrons ($M \cong \frac{1}{2000}$, $w = 2$, and $\rho = 3150$ for

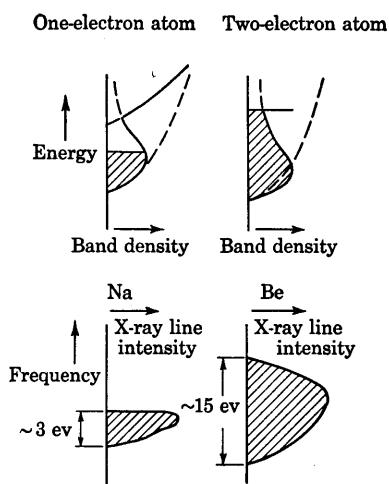


FIG. 7.24. Illustrative band occupancy and X-ray-line contour comparisons.

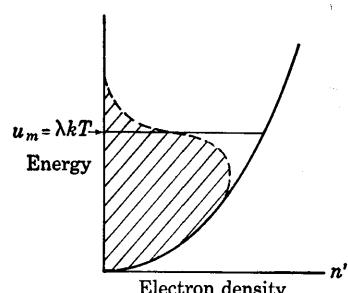


FIG. 7.25. Schematic of electron density for electron gas in a metal.

one electron per atom in a copper crystal, and $T = 2000^{\circ}\text{abs}$), $e^{-\lambda} = 4.4 \times 10^{-3}$. The perfect-gas approximation which was seen to hold for large negative values of λ is evidently entirely inadequate. $e^{-\lambda}$ is so small that to a satisfactory approximation for many purposes

$$\frac{1}{e^{-\lambda+x^2} + 1} \approx \begin{cases} 0 & x > \sqrt{\lambda} \\ 1 & x < \sqrt{\lambda} \end{cases} \quad (7.32)$$

Under these circumstances

$$n' = 8\pi \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \int_0^{\sqrt{\lambda}} x^2 dx = \frac{8\pi}{3} \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \lambda^{\frac{3}{2}} \quad (7.33)$$

$$\bar{u} = 8\pi \frac{kT}{n'} \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \int_0^{\sqrt{\lambda}} x^4 dx = \frac{8\pi}{5} \left(\frac{2mkT}{h^2} \right)^{\frac{3}{2}} \lambda^{\frac{5}{2}} \frac{kT}{n'}$$

Eliminating λ and using Eq. (6.28),

$$\bar{u} = \frac{3^{\frac{5}{2}}}{20} \left(\frac{n'}{\pi} \right)^{\frac{3}{2}} \frac{h^2}{2m} = \frac{3}{2} p V = \frac{3}{2} \frac{p}{n'}$$

and

$$p = \frac{1}{10} \frac{h^2}{2m} \left(\frac{3}{\pi} \right)^{\frac{3}{2}} n'^{\frac{5}{2}} \quad (7.34)$$

¹ J. A. Becker, *Rev. Mod. Phys.*, **7**, 95 (1935); R. O. Jenkins, *Repts. Progr. in Phys.*, **11**, 177 (1942); H. Friedenstein, S. L. Martin, and G. L. Mundan, *Repts. Progr. in Phys.*, **11**, 298 (1948); C. Herring and M. H. Nichols, *Rev. Mod. Phys.*, **21**, 185 (1949).

To this approximation \bar{u} is independent of the temperature so the specific heat of such an electron gas vanishes and the electrons contribute nothing to the lattice specific heat previously calculated. Also pV^* is a constant which is the same as the adiabatic equation of a perfect monatomic gas. The density of electrons on an energy scale is $4V\pi(2m^*/h^2)^{1/2}\sqrt{\bar{u}}$ up to u_m , as in the previous section, and zero from there on as shown in the solid line of Fig. 7.25.

To a more accurate approximation for which the references should be consulted the mean value of the energy of an electron is given by

$$\bar{u} = \frac{3u_m}{5} \left[1 + \frac{5\pi^2}{12} \left(\frac{kT}{u_m} \right)^2 \right]$$

and the atomic heat $(\partial\bar{u}/\partial T)$ is $\frac{k\pi^2}{2} \frac{kT}{u_m}$. Since \bar{u} or u_m is of the order of several electron volts, this contribution to the specific heat is very small. However, it is an appreciable fraction of the total specific heat for a crystal at very low temperatures and is responsible for the linear dependence of C_a on T mentioned in the preceding section.

In the liberation of electrons by the external photoelectric effect or thermionic emission there is a very considerable potential barrier, at least partly due to the image force of an electron on leaving the crystal, that must be overcome before any current of electrons from the surface is observed. Thus but a very small fraction of the electrons in the band, those of the very highest energy, are the ones observed to leave near the photoelectric threshold or as the crystal temperature is raised. The distribution function for a finite temperature is indicated by the dashed curve in Fig. 7.26, and the present concern is with the very small peak of most energetic electrons only. The potential affecting an electron near the surface of a metal is made up primarily of the image potential and that of any external field such as one tending to remove an electron as illustrated schematically in Fig. 7.26. Writing the effective potential for an electron leaving the crystal as that of the image potential plus the applied field potential,

$$\varphi = -\frac{e^2}{4\pi\epsilon_0(4x)} - eE_x x$$

The maximum, $\partial\varphi/\partial x = 0$, occurs at $x_m^2 = e/16\pi\epsilon_0 E$, or writing φ_0 for the height of the image-force barrier alone above u_m the potential barrier actually to be surmounted by an electron initially of energy u_m is

$$\varphi_m = \varphi_0 - \left(\frac{Ee^3}{4\pi\epsilon_0} \right)^{1/2}$$

The electron current per unit area in the x direction within the electrons in the crystal band is the product of the number of electrons per

unit volume in the elementary momentum range $dp_x dp_y dp_z$ times the charge and the x component of the velocity integrated over all values of p_y and p_z and over those values of p_x for which the electron could pass over the barrier with a value of v_x greater than zero. The number of

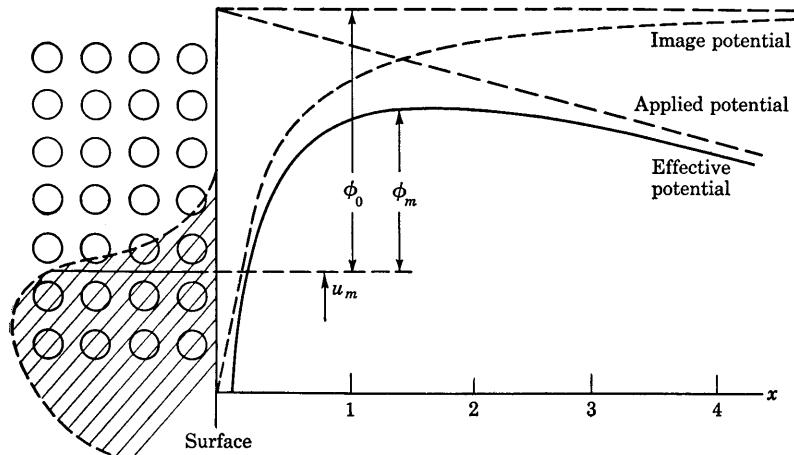


FIG. 7.26. Potential relations for an electron leaving a metal surface.

electrons per unit volume with momentum components in the ranges dp_x , dp_y , and dp_z is

$$n'(p) dp_x dp_y dp_z = \frac{2}{e^{-\lambda} e^{u(p)/kT} + 1} \frac{dp_x dp_y dp_z}{h^3}$$

from Eq. (6.18F); hence the current per unit area is

$$I' = e \int_{p_z} \int_{p_y} \int_{p_x} \frac{2v_x}{e^{-\lambda} e^{u(p)/kT} + 1} \frac{dp_x dp_y dp_z}{h^3}$$

To surmount the barrier φ_m , the component of velocity or momentum in the x direction must be adequate, that is, $(1/2m)p_x^2 > u_m + \varphi_m$, so the limits of integration are

$$I' = \frac{e}{mh^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{p_x^*}^{\infty} \frac{2p_x dp_x dp_y dp_z}{e^{-\lambda} e^{u/p/kT} + 1}$$

where $[(1/2m)p_x^2]^* = u_m + \varphi_m$. From the limiting approximation of Eq. (7.29), $\lambda = u_m/kT$, and the total exponent in the denominator is $(1/kT)[u_m - \frac{1}{2}m(p_x^2 + p_y^2 + p_z^2)]$. The p_y and p_z integrations can be done immediately by the substitution $(p_y^2 + p_z^2)/2mkT = \xi^2$, in which case the double-integration element is $dp_y dp_z = 4\pi mkT\xi d\xi$ and the integration limits are 0 to ∞ .

$$\begin{aligned} I' &= \frac{2\pi kTe}{h^3} \int_0^\infty \int_{p_x*}^\infty \frac{2p_x dp_x 2\xi d\xi}{e^{-(1/kT)[u_m - (1/2m)p_x^2]} e^{\xi^2} + 1} \\ &= \frac{2\pi kTe}{h^3} \int_{p_x*}^\infty 2p_x dp_x \ln (1 + e^{(1/kT)[u_m - (1/2m)p_x^2]}) \end{aligned}$$

Making the substitution $y = (1/2m)p_x^2 - u_m - \varphi_m$, the lower limit is zero and the upper still ∞ , the exponent is $-(y + \varphi_m)/kT$,

$$2p_x dp_x = 2m dy$$

and

$$I' = \frac{4\pi mkTe}{h^3} \int_0^\infty \ln (1 + e^{-(y+\varphi_m)/kT}) dy \quad (7.35)$$

This cannot be integrated in a closed form as it stands, but as φ_m is found experimentally to be of the order of 1 to 10 ev the negative exponent is large even for high temperatures and the logarithm may be approximated by the exponential term, in which case the integration may be performed immediately to yield

$$I' = \frac{4\pi m(kT)^2 e}{h^3} e^{-\varphi_m/kT}$$

or

$$I' = A T^2 e^{-b/T} \quad (7.36)$$

where

$$A = \frac{4\pi mk^2 e}{h^3}, \quad b = \frac{\varphi_m}{k}$$

Equation 7.36 is known as the *Richardson-Dushman equation* for thermionic emission, and its general form has been verified over a very great range of the variable T . The constant A has the value 1.2×10^6 amp m⁻² as may be seen by substituting the atomic constants, and this is the proper order of magnitude to accord with observation. Both A and b are quite sensitive to the crystal surface conditions, and a much more detailed analysis would be necessary to attempt to account for the variation of these values with variation in surface character. The dependence of φ_m on the applied field E given earlier has been well verified and is known as the *Schottky effect* or the *field emission effect*. Representative values of φ_0 are given in table 7.8.

In the surface photoelectric effect a photon is absorbed by an electron in the band, and provided its resulting x component of momentum is sufficiently great to surmount the surface barrier it will leave the metal and a photoelectric current will be observed. It is evident that the theory of this effect parallels very closely the foregoing except that the energy of an electron under consideration is $u + h\nu$ so that the electrons

that can escape are those for which

$$\left(\frac{1}{2m} p_x^2\right)^* \geq u_m + \varphi_m - h\nu$$

The photoelectric current may be worked out as a universal function of $(\varphi_m - h\nu)/kT$ from Eq. (7.35); for the details the references should be consulted. φ_m is the minimum barrier to be overcome for the principal upper occupancy levels at low temperature and thus corresponds to the standard threshold energy $h\nu_0$, where ν_0 is the minimum photon frequency that will liberate an electron at $T = 0$. The curves are in excellent agreement with experiment and thus amply justify the correctness of the various hypotheses involved. The comparison of the parameters φ_m and

TABLE 7.8. METALLIC WORK FUNCTIONS

Metal	φ_0, ev
Cs	1.8
Ca	3.2
Th	3.4
Ta	4.1
Mo	4.3
W	4.5
Pt	5.0

$h\nu_0$ from thermionic and photoelectric effects yields very satisfactory agreement for different materials if due allowance is made for surface effects and the results are reduced to the same temperature.

PROBLEMS

- 7.1. One method of detecting the presence of a molecular beam is by observing the pressure that is built up within a chamber into which the beam may be directed through a small opening generally in the form of a cylindrical tube coaxial with the direction of the beam. The equilibrium is brought about in the following way: Atoms from the oven, where the pressure is p_1 and the temperature T_1 , emerge from an opening of area a_1 and are diaphragmed into a narrow pencil of solid angle $d\omega$ which is so small that all the atoms are shot directly through the tube into the interior of the chamber. The atoms in the chamber, where the pressure is p_2 and the temperature T_2 , tend to leak back out of the cylindrical opening through which the beam enters. Assume that the area of the opening is a_2 and that a fraction f of the atoms in the chamber, determined by the nature of the cylindrical openings, which strike the area a_2 from all internal directions emerge. Show that the pressure of atoms in the chamber is given in equilibrium by

$$p_2 = p_1 \sqrt{\frac{T_2}{T_1}} \frac{a_1}{a_2} \frac{d\omega}{2\pi f}$$

- 7.2. Stern proposed a molecular-beam experiment in which the gravitational force on the horizontal beam of atoms would be balanced by the magnetic force and the magnetic moment measured in terms of the mass, magnetic gradient, and gravitational

acceleration. Show that a magnetic-induction gradient of 4.03×10^{-2} weber m $^{-3}$ would just balance the gravitational force in the case of a properly oriented component of a beam of sodium atoms. Show that, if a beam of atoms of mass m emerges in a narrow, collimated horizontal pencil from an oven at a temperature T and traverses a distance l in a region where the gravitational acceleration is g , the maximum in the distribution of atoms striking unit area per second along the vertical coordinate z occurs at a distance $mgl^2/12kT$ below the axis of the beam emerging from the oven.

7.3. A narrow beam of atoms of mass m emerges from a hole in the walls of a furnace maintained at the temperature T and impinges on a crystal surface from which diffraction is observed. Show that the effective wavelength of the atoms corresponding to the maximum of the curve representing the distribution of number of atoms striking the crystal per unit time as a function of wavelength occurs at

$$h/\sqrt{5mkT} = 1.959 \times 10^{-9} (MT)^{-\frac{1}{2}},$$

where M is the atomic weight of the atoms in the beam.

7.4. Show by the use of Eqs. (7.8') and (7.9) that the mean value of v is $2v_m/\sqrt{\pi}$, the mean value of v^2 is $3v_m^2/2$, the mean value of v_x over positive x is $v_m/(2\sqrt{\pi})$, and the mean value of v_x^2 over all x is $v_m^2/2$.

7.5. Assuming all atoms striking a surface stick to it in a pattern of closely packed spheres, show that the time in which a monatomic layer of gas atoms of radius r would be formed on the surface is $\sqrt{\pi/3}(v_m n' r^2)^{-1}$. Calculate this time for helium at 10^{-6} atm pressure and 300°abs assuming the atomic radius to be 10^{-10} m.

7.6. From the definitions $\mu = \frac{1}{2} \sum_j e_j \mathbf{r} \times \mathbf{v}$, and $P = \sum_j e_j l_i$, and taking uniform electric and magnetic conditions such that

$$\varphi = -E_z z, \quad A_x = -\frac{1}{2}B_z y, \quad A_y = \frac{1}{2}B_z x, \quad A_z = 0$$

in terms of cartesian coordinates, show that the magnetic and electric moments of a system described by the Hamiltonian $H = (1/2m)(p - eA)^2 + e\varphi$ are given by $\mu = -\partial H / \partial B$ and $P = -\partial H / \partial E$ respectively.

7.7. The atomic magnetic energy $-\mu B$ for a material in which the atoms or molecules have a mean moment μ in the presence of an effective orienting magnetic induction B can be incorporated in the general incremental equation of thermodynamics for the energy in a manner analogous to that for the mechanical-energy term pV . Show that, on including the magnetic energy, Eq. (1.8) for the internal energy becomes

$$dU = T dS - p dV + B d\mu$$

and derive the differential expressions for the other thermodynamic potentials. Consider a perfect paramagnetic gas occupying a region of nonuniform magnetic induction in thermal equilibrium at a high temperature. Show by the use of the appropriate thermodynamic potential that the variation of density ($n' = 1/V$) with magnetic induction is given by

$$\frac{1}{n'} \frac{dn'}{dB} = \frac{J(J+1)}{3} \left(\frac{\mu_0 g}{kT} \right)^2 B$$

7.8. Using the expression for the increment of internal energy given in Prob. 7.7 under conditions of constant volume and considering μ to be a function of B and T only, show that the atomic heat in the presence of a constant magnetic induction is greater than in the absence of a magnetic induction by the term $-B(\partial\mu/\partial T)_B$.

Show that for a paramagnetic gas this is

$$k\beta^2 \left\{ \frac{1}{4} \operatorname{csch}^2 \frac{\beta}{2} - (J + \frac{1}{2})^2 \operatorname{csch}^2 [(J + \frac{1}{2})\beta] \right\}$$

where $\beta = \mu_0 g B / kT$. Show that at high temperatures this reduces in units of k to $[J(J+1)/3]\beta^2$. Calculate this increment for atomic hydrogen ($J = \frac{1}{2}$ and $g = 2$) for the conditions that $B = 1$ weber m⁻² and $T = 300^\circ\text{abs}$.

7.9. Using the concepts of the Thomas-Fermi atom from Sec. 6.5 and Eq. (7.5), show that the atomic diamagnetic susceptibility for a monatomic gas is given by

$$-1.029 \times 10^{-29} Z^{\frac{1}{3}} \int_0^{\infty} X^{\frac{8}{3}} x^{\frac{5}{3}} dx \quad \text{amp weber}^{-1} \text{m}^4$$

7.10. Use the wave function for the ground state of atomic hydrogen from Sec. 3.4 to calculate the diamagnetic moment, and assuming that all the atoms remain in the ground state, show that the susceptibility of atomic hydrogen gas is given by

$$\chi_m = n' \left(\frac{\mu_0^2}{kT} - \frac{e^2 a_0^2}{2m_0} \right)$$

and that this vanishes at a temperature such that kT equals half the energy of the ground state of the atom.

7.11. Consider a perfect gas in the temperature range to which the classical statistics is applicable consisting of molecules with an electric polarizability α and a permanent electric dipole moment P_0 . Assuming the classical approximation in which the angle between a molecular moment and the applied electric field E is a continuous variable, show that the distribution function in electrical energy is $(4\pi \sinh x)/x$, where $x = P_0 E/kT$. Show that the electric susceptibility of the gas is given by

$$n'[\alpha + (P_0/E)(-1/x + \coth x)]$$

which reduces to $n'(\alpha + P_0^2/3kT)$ for high temperatures. n' is the number of molecules per unit volume.

7.12. Measurements of the electric susceptibility of HCl vapor by Zahn have shown that the dependence of the susceptibility on the temperature over a considerable range is given by the expression in Prob. 7.11. Zahn's measurements were made at constant pressure (1 atm), and he found that the electric susceptibility in practical units was 6.598×10^{-14} farad m⁻¹ at 201.4°abs and 1.0465×10^{-14} farad m⁻¹ at 588.8°abs . Show that P_0 for the HCl molecule is equal to 0.400 in units of the product of the electron charge and Bohr radius. Show that α for the HCl molecule is given by 0.697 in units of the square of the product of the electron charge and Bohr radius per electron volt of energy.

7.13. Consider a homogeneous chemical reaction taking place between two gases the molecules of which are of such a nature that the reaction does not take place at all collisions for which the energy in the center-of-gravity system exceeds the critical value u_c as in the text but that the cross section for the reaction is greatest for a particular relative energy u_c and, in particular, can be expressed as $\sigma'' e^{-(1-u/u_c)^2/(4\delta)^2}$. Show that the reaction rate under the general simplifying assumptions of the text is then given by

$$\frac{n'_1 n'_2 \sigma'' u_c \delta}{2 \sqrt{k T M_r}} e^{-(u_c/kT)(1-u_c \delta^2/8kT)}$$

if δ is very small compared to unity but $u_c \delta$ is not very large compared to kT .

7.14. Assume that the energy states available to a molecule are equally spaced and N in number, as would be the case for the magnetic states if the spin-angular-momentum components could take $N = 2j + 1$ orientations, and that they are separated by an interval du . Show that the statistical distribution function Z is $Z = e^{-(N-1)x} \sinh(Nx)/\sinh x$, where $x = du/2kt$, that the mean energy \bar{u} is $\bar{u} = (du/2)[(N - 1) - N \coth Nx + \coth x]$, and that the molecular heat C_a is $C_a = k[x^2 \cosh^2 x - (Nx)^2 \cosh^2 Nx]$. Show graphically how C_a varies with x and with the parameter N .

7.15. Using van der Waals' equation and the values given in the text for the constants in terms of atomic quantities, show that the values of the volume, temperature, and pressure at the critical point are given by

$$V_c = \frac{4\pi Na_0^3}{Z}, \quad T_c = \frac{64E_0^2 Z^{\frac{5}{2}}}{81E_i k}, \quad p_c = \frac{2E_0^2 Z^{\frac{3}{2}}}{27\pi E_i a_0^3}$$

Compare the calculated values of T_c and the product $V_c p_c$ with the experimental data for the gases listed in the following table.

Gas	T_c , deg C	p_c , atm	ρ_c , g cm $^{-3}$	E_i , ev
He	-267.9	2.26	0.693	24.46
Ne	-228.7	25.9	0.484	21.47
A	-122.0	48	0.531	15.68
Kr	-63.0	54	0.78	13.93
Xe	16.6	58.2	1.156	12.08

7.16. A gas reaction is represented by the equation $\sum_j \nu_j A_j = 0$, where ν_j is the number of molecules of type A_j taking part in a unit reaction. From the condition that the free energy is a minimum as a function of concentration at equilibrium, $\sum_j \Psi_j \delta n_j = 0$, where δn_j is the proportional change in the component A_j , show that the equilibrium condition for the reaction between perfect gases is

$$\sum_j \nu_j \left(\ln \frac{n_j}{V} - 1 - \ln \omega_j - \frac{3}{2} \ln \frac{2\pi m_j kT}{h^2} - \ln Z'_j \right) = 0$$

where Z'_j represents all the factors in the partition function other than the monatomic translation factor. If Z'_j for each type of molecule is written as the product of the distribution function Z''_j relative to the lowest, or fiducial, energy level of that molecule and $e^{-u_j/kT}$, where u_j is the energy of this fiducial level on a common scale for all the reacting molecules, show that the equilibrium condition in terms of the partial pressures $p_j = n_j kT/V$ of the component gases is

$$\prod_j p_j^{\nu_j} = \prod_j (\omega_j Z''_j e)^{\nu_j} \left(\frac{2\pi m_j}{h^2} \right)^{3\nu_j/2} (kT)^{5\nu_j/2} e^{-\nu_j u_j/kT}$$

7.17. Consider the chemical reaction in which molecules of hydrogen and deuterium are in equilibrium with the mixed molecule HD, that is, $H_2 + D_2 \rightleftharpoons 2HD$ on the assumption that the electronic structures of the H and D atoms are identical and that the molecules differ only through the difference in mass between the H and D atoms. Considering simply the translational, vibrational, and rotational partition functions,

which are the only ones of concern, and making due allowance for the different statistical weights w of the different molecular types show that for temperatures such that the rotational motion is fully developed but considerably lower than $\hbar\omega_c/k$

$$\frac{N_{\text{HD}}^2}{N_{\text{H}_2} N_{\text{D}_2}} = 4 \left(\frac{9}{8}\right)^{\frac{1}{2}} e^{-(\hbar/2kT)(2\omega_c \text{HD} - \omega_c \text{H}_2 - \omega_c \text{D}_2)} \\ \cong 4.25 e^{-86.7/T}$$

For $T \gg \hbar\omega_c/k$, show that the ratio approaches 4.

7.18. Calcium vapor in the sun undergoes ionization as a reversible chemical reaction, $\text{Ca} \rightleftharpoons \text{Ca}^+ + e^-$. If X is the fraction of calcium that is ionized, show that

$$\frac{1 - X^2}{X^2} = \frac{p_t}{4\epsilon_0 k T} \left(\frac{\hbar^2}{2\pi m_0 k T} \right)^{\frac{1}{2}} e^{eV_i/kT}$$

where p_t is the total pressure (sum of the partial pressures, $p_{\text{Ca}} + p_{\text{Ca}^+} + p_{e^-}$), m_0 is the electron mass, and V_i is the ionization potential of calcium. The spin angular momenta are zero for calcium and $\frac{1}{2}$ for the electron and the ion. The equation is known as *Saha's equation* for the ionization in stellar atmospheres.

7.19. Under certain simplifying assumptions (Prob. 4.18) the interaction of the aligned atomic magnetic moments distributed uniformly throughout a crystal is such that the magnetic induction B effective at the position of one atom due to all the rest is $-2(3\epsilon_0 c^2)^{-1}$ times the magnetic moment per unit volume. Using this result and Eq. (7.7) for low temperatures, calculate the temperature at which the magnetic energy of orientation of an atom for which $J = \frac{1}{2}$ and $g = 2$ in a crystal would equal the thermal energy kT . Use the density ($\rho = 8900 \text{ kgm}^{-3}$) and atomic weight ($M = 58.69$) of nickel as representative in the calculation. What does this result imply as to the adequacy of magnetic forces in maintaining atomic-moment alignment at ordinary temperatures?

7.20. Using Eq. (7.33), calculate the energy of the top of the filled band $u_m = \lambda kT$ and the average electron energy \bar{u} for copper.

7.21. Assume the simple model of a paramagnetic crystal in which the average magnetic moment per atom in the direction of an applied B is given by Eq. (7.7'') and the specific heat is given by Eq. (7.25'D) for $g\mu_0 B \ll kT \ll h\nu_m$. Show that for a process in which the crystal is thermally isolated ($\delta Q = T ds = du - B d\mu_p = 0$), the variation of T and B is given by

$$\frac{dT}{dB} = \frac{T}{B} \left[\frac{27(\pi^2/90)^2}{J(J+1)} \left(\frac{kT}{h\nu_m} \right)^3 \left(\frac{kT}{g\mu_0 B} \right)^2 + 1 \right]^{-1}$$

7.22. Using Eq. (7.31) for the magnetization of a ferromagnetic material above the Curie point, and taking $J = \frac{1}{2}$ and $g = 2$, show that the atomic susceptibility for $T > \theta$ is given by

$$\chi_\alpha = \frac{a\mu_0^2}{k(T - \theta)}$$

where a is the number of Bohr magnetons associated with each atom. Note that the magnetic moments per unit volume of the text can be written as $I = n'\mu$ and $I_0 = an'\mu_0$, where n' is the number of atoms per unit volume.

7.23. For a ferromagnetic substance the value of B in the incremental equation for the free energy must be replaced by $B + CI$ of the text. Neglecting any change in the volume, show that if the magnetic moment of a thermally insulated ferromagnetic substance is changed by an amount $d\mu$ the temperature change induced is

given by

$$dT = \frac{B + \mu(\theta k/a\mu_0^2)}{C_a} d\mu$$

where C_a is the atomic heat.

7.24. Using the results of the preceding problems show that there is a term in the expression for the specific heat per atom of a ferromagnetic substance due to the internal forces tending to orient the magnetic moments which is given by $-\mu \frac{\theta k}{a\mu_0^2} \frac{\partial \mu}{\partial T}$.

By evaluating μ^2 from Eq. (7.30) just below the Curie point, show that there is a discontinuity at that point in the specific heat per atom as a function of the temperature which is given by $\Delta C_a = \frac{3}{2}ak$.

GENERAL REFERENCES

- Bates, L. F.: "Modern Magnetism," Cambridge University Press, London, 1950.
- Bozorth, R. M.: "Ferromagnetism," D. Van Nostrand Company, Inc., New York, 1951.
- Chapman, S., and T. G. Cowling: "Non-uniform Gases," Cambridge University Press, London, 1939.
- Dushman, S.: "High Vacuum," John Wiley & Sons, Inc., New York, 1949.
- Fowler, R. H.: "Statistical Mechanics," Cambridge University Press, London, 1929.
- Frenkel, J.: "Kinetic Theory of Liquids," Oxford University Press, London, 1946.
- Glasstone, S., K. J. Laidler, and H. Eyring: "Theory of Rate Processes," McGraw-Hill Book Company, Inc., New York, 1941.
- Gurney, R. W.: "Introduction to Statistical Mechanics," McGraw-Hill Book Company, Inc., New York, 1949.
- Hughes, A. L., and L. A. DuBridge: "Photoelectric Phenomena," McGraw-Hill Book Company, Inc., New York, 1932.
- Kennard, E. H.: "Kinetic Theory of Gases," McGraw-Hill Book Company, Inc., New York, 1938.
- London, F.: "Superfluids," John Wiley & Sons, Inc., New York, 1950.
- Mayer, J., and M. G. Mayer: "Statistical Mechanics," John Wiley & Sons, Inc., New York, 1940.
- Mott, N. F., and R. W. Gurney: "Electronic Processes in Ionic Crystals," Oxford University Press, London, 1940.
- Mott, N. F., and H. Jones: "The Theory of Properties of Metals and Alloys," Oxford University Press, London, 1936.
- Riemann, A. L.: "Thermionic Emission," John Wiley & Sons, Inc., New York, 1934.
- Schroedinger, E.: "Statistical Thermodynamics," Cambridge University Press, London, 1946.
- Seitz, F.: "The Modern Theory of Solids," McGraw-Hill Book Company, Inc., New York, 1940.
- Seitz, F.: "The Physics of Metals," McGraw-Hill Book Company, Inc., New York, 1943.
- Slater, J. C.: Electron Theory of Solids, *Am. J. Phys.*, **19**, 368 (1951).
- Slater, J. C.: "Quantum Theory of Matter," McGraw-Hill Book Company, Inc., New York, 1951.
- Van Vleck, J. H.: "Electric and Magnetic Susceptibilities," Oxford University Press, London, 1946.

APPENDIX

RELATIONSHIPS BETWEEN SYSTEMS OF UNITS COMMONLY EMPLOYED IN ATOMIC PHYSICS

Entity	mks	Electrostatic	Atomic
Length(l).....	1 m	10^2 cm	$1.890 \times 10^{10} a_0$
Mass(m).....	1 kg	10^3 g	$1.098 \times 10^{30} m_0$
Time(t).....	1 sec	1 sec	$4.134 \times 10^{16} t_0$
Frequency(ν).....	1 sec^{-1}	1 sec^{-1}	$2.419 \times 10^{-17} t_0^{-1}$
Velocity(v).....	1 m sec^{-1}	10^2 cm sec^{-1}	$4.571 \times 10^{-7} (a_0/t_0 = \alpha c)$
Density(ρ).....	1 kg m^{-3}	$10^{-3} \text{ g cm}^{-3}$	$0.1627 m_0/a_0^3$
Angular momentum(L).....	$1 \text{ kg m}^2 \text{ sec}^{-1}$	$10^7 \text{ g cm}^2 \text{ sec}^{-1}$	$9.846 \times 10^{33} (\hbar = m_0 a_0^2/t_0)$
Force(F).....	1 newton	10^5 dynes	$1.214 \times 10^7 m_0 a_0/t_0^2$
Energy(E).....	1 joule	10^7 ergs	$2.295 \times 10^{17} (\hbar/t_0 = m_0 c^2 \alpha^2 = 2 E_0)$
Pressure(p).....	1 newton m^{-2}	10 dynes cm^{-2}	$3.399 \times 10^{-14} (m_0 t_0^{-2} a_0^{-2})$
Charge(q).....	1 coulomb	2.998×10^9 esu	$6.243 \times 10^{19} e$
Current(i).....	1 amp	2.998×10^9 esu	$151.2 e/t_0$
Dipole moment(P).....	1 coulomb-m	2.998×10^{11} esu	$1.180 \times 10^{29} e a_0$
Potential(V).....	1 volt	3.335×10^{-3} esu	$3.677 \times 10^{-8} \hbar (e t_0)^{-1}$
Field(E).....	1 volt m^{-1}	3.335×10^{-6} esu	$1.946 \times 10^{-12} \hbar (e a_0 t_0)^{-1}$
Electric			
susceptibility(χ_e)....	1 coulomb volt $^{-1} \text{m}^{-1}$	8.987×10^9 esu	$8.983 \times 10^9 \hbar (t_0 a_0)^{-1}$
Capacity(C).....	1 farad	8.987×10^{11} esu	$1.697 \times 10^{20} e^2 t_0/\hbar$
Resistance(R).....	1 ohm	1.113×10^{-12} esu	$2.490 \times 10^{-4} \hbar e^{-2}$
Inductance(L).....	1 henry	1.113×10^{-12} esu	$1.007 \times 10^{13} \hbar t_0 e^{-2}$
Flux(ϕ).....	1 weber	3.335×10^{-3} esu	$1.520 \times 10^{15} \hbar (e t_0)^{-1}$
Magnetic induction(B)..	1 weber m^{-2}	3.335×10^{-7} esu	$4.253 \times 10^{-8} m_0 (e t_0)^{-1}$
Magnetic moment(μ)..	1 amp-m 2	2.998×10^{13} esu	$5.395 \times 10^{22} e \hbar / m_0$
Magnetic			
susceptibility(χ_m)....	1 amp weber $^{-1} \text{m}$	8.996×10^{13} esu	$1.493 \times 10^{-3} e^2 a_0 (\hbar t_0)^{-1}$

a_0 = Bohr radius, m_0 = electron mass, t_0 = Bohr time, e = electron charge, \hbar = Planck's constant.
 α = fine structure constant, E_0 = Bohr energy.

PRINCIPAL FUNDAMENTAL ATOMIC CONSTANTS

Quantity	Absolute practical units physical scale of atomic mass units ($0^{16} = 16$)	Stand- ard error in parts per million
Velocity of light.....	$c = 2.997929 \times 10^8 \text{ m sec}^{-1}$	3
Avogadro's number.....	$N = 6.02474 \times 10^{23} \text{ molecules mole}^{-1}$	60
Atomic weight mass unit.....	$M_1 = 1/N = 1.65982 \times 10^{-27} \text{ kg}$	60
Faraday.....	$F = 96520.1 \text{ coulomb g equivalent}^{-1}$	26
Electronic charge.....	$e = FM_1 = 1.60207 \times 10^{-19} \text{ coulomb}$	43
Specific electronic charge.....	$e/m_0 = 1.75888 \times 10^{11} \text{ coulomb kg}^{-1}$	28
Electron rest mass.....	$m_0 = 9.1085 \times 10^{-31} \text{ kg}$	66
Proton-electron mass ratio.....	$M_p/m_0 = 1836.3$	21
Proton atomic weight.....	$M_p = 1.007593 \text{ amu}$	3
Electron atomic weight.....	$M_e = Nm_0 = 5.48760 \times 10^{-4} \text{ amu}$	23
Planck's constant.....	$h = 6.6252 \times 10^{-34} \text{ joule-sec}$	75
Bohr magneton.....	$\mu_0 = e\hbar/2m_0 = 0.92732 \times 10^{-23} \text{ amp m}^2$	64
Reciprocal fine structure constant.....	$1/\alpha = 4\pi\epsilon_0 c\hbar/e^2 = 137.0377$	12
Classical electron radius.....	$r_0 = e^2/(4\pi\epsilon_0 m_0 c^2) = 2.81784 \times 10^{-15} \text{ m}$	36
Compton wavelength.....	$\lambda_c = 2\pi r_0/\alpha = h/m_0 c = 2.42625 \times 10^{-12} \text{ m}$	11
First Bohr radius.....	$a_0 = r_0/\alpha^2 = 4\pi\epsilon_0 \hbar^2/m_0 e^2 = 5.29171 \times 10^{-11} \text{ m}$	11
Rydberg constant.....	$R_\infty = \alpha^2/2\lambda_c = 10973730.9 \text{ m}^{-1}$	0.1
Boltzmann's constant.....	$k = 1.38042 \times 10^{-23} \text{ joule } (\text{°C})^{-1}$	72
Gas constant.....	$R_0 = kN = 8.31662 \text{ joule mole}^{-1} (\text{°C})^{-1}$	46
Volume of 1 kilomole at NTP.....	$V_0 = 22.4207 \text{ m}^3$	27
Ice point.....	$T_0 = 273.16^\circ\text{abs}$	37
Mechanical equivalent of heat.....	$J_{15} = 4.1855 \text{ joule cal}^{-1}$	100
Gravitational constant.....	$G = 6.670 \times 10^{-11} \text{ newtons-m}^2 \text{ kg}^{-2}$	750
1 ev equivalent Wavelength.....	$\lambda_0 = 1.23978 \times 10^{-6} \text{ m}$	40
Frequency.....	$\nu_0 = 2.41579 \times 10^{14} \text{ sec}^{-1}$	40
Wave number.....	$\tilde{\nu}_0 = 806598 \text{ m}^{-1}$	40
Temperature.....	$T_0 = 11,605.6^\circ\text{C}$	50
Energy.....	$u_0 = 1.60207 \times 10^{-19} \text{ joule}$ $= 23.060 \text{ kilocal mole}^{-1}$	40
1 electron mass.....	$m_0 c^2 = 0.510984 \times 10^6 \text{ ev}$	31
1 amu.....	$M_1 c^2 = 931.162 \times 10^6 \text{ ev}$	20
Permittivity of space.....	$\epsilon_0 = (4\pi c^2 \times 10^{-7})^{-1}$ $= 8.8542 \times 10^{-12} \text{ farad m}^{-1}$	10

SOURCE: J. W. M. DuMond and E. R. Cohen, *Rev. Mod. Phys.*, **25**, 691 (1953) and R. T. Birge, *Rev. Mod. Phys.*, **13**, 233 (1941).

INDEX

- Absorption, of electromagnetic radiation, 143, 147
magnetic resonance, 80
microwave, 114
quadrupole, 148
- Acceleration, 6
- Acetylene, 249
- Activation energy, 337
- Adams, N. I., 39
- Allen, S. J. M., 151
- Allison, S. K., 93, 159, 279
- Allowed transitions, 98
- Alpha matrices, 171
- α -particle scattering, 100
atomic number by, 106
- Amagat units, 348
- ämpère, A. M., 23, 24
- Anderson, C. D., 167
- Anderson, H. L., 82
- Angstrom bands of CO, 269
- Angular momentum, 9, 128
of atomic nuclei, 79
conservation of, 10, 173
electronic, 241, 252
intrinsic, 75
of particles, 70
- Angular-momentum quantization, 126
- Angular wave functions, 127
- Antiferroelectric moments, 372
- Antiferromagnetic moments, 372
- Antisymmetric wave functions, 199
- Arnold, W. R., 82
- Aslakson, C. I., 115
- Associated Legendre functions, 126
- Aston, F. W., 93
- Atmospheres, law of, 42
- Atomic-beam apparatus, 111
- Atomic-beam zero-moment method, 192
- Atomic constants, 390
- Atomic heat, 325
- Atomic mass unit, 56
- Atomic masses, 42, 43, 89
- Atomic nuclei, 79
angular momentum of, 79
- Atomic number, by α -particle scattering, 106
- Atomic number, by X-ray spectra, 106
- Atomic physics, 5
- Atomic structure, 95 ff.
- Atomic theory, classical foundations of, 3 ff.
- Atomic transitions, higher-energy, 115
low-energy, 110
- Atomic units, 389
- Atomic weight, 56
chemical scale of, 49
- Atomic-weight mass unit, 390
- Atomicity, evidence for, 40
of matter, 4
- Attenuation coefficient, 114
- Auger effect, 120
- Autoionization, 120
- Avogadro, 41
- Avogadro's number, 42, 390
- Bacher, R. F., 159, 207, 222
- Bainbridge, K. T., 56, 57, 89, 94
- Balmer, E. J., 95
- Band, electron, 368
- Band head, 270
- Band theory, electron, 364
- Barkla, C. G., 101
- Barlow, H. S., 151
- Barnes, R. B., 296
- Barnett, S. J., 373
- Barrett, C. S., 229
- Bates, L. F., 387
- Bearden, J. A., 48, 64
- Beattie, J. A., 348
- Becker, G. E., 190
- Becker, J. A., 378
- Bennett, W. H., 56
- Bergmann, P., 39
- Bergstrand, E., 116
- Beringer, B., 113
- Bernoulli, D., 40
- Bewilogna, L., 229
- Binding energy, 202, 242
- Birge, R. T., 44, 48, 55, 306
- Bitter, F., 82
- Black-body radiation, 304

- Black-body radiation law, 306
 Blackman, M., 364
 Bleany, B., 112
 Bloch, F., 81, 82, 84
 Bloembergen, N., 80, 93
 Bohr, N., 98
 Bohr frequency relationship, 98
 Bohr magneton, 82, 390
 Bohr-orbit radius, 133
 Bohr-Rutherford hydrogen atom, 121
 Boiling point, 349
 Boltzmann distribution law, 288
 Boltzmann's constant, 43, 390
 Born, M., 69, 93, 233, 359
 Bose-Einstein distribution, 301
 Bose-Einstein gas, 355, 357
 Bose-Einstein statistics, 301
 Boyce, J. C., 118
 Boyle, R., 41
 Boyle temperature, 348
 Bozorth, R. M., 372, 373, 387
 Brackett, F. S., 96
 Bragg law, 47
 Breit, G., 190, 192
 Bremsstrahlung, 153
 Bridgman, P. W., 39
 Brillouin, L., 315
 Brillouin function, 332
 Brown, R., 43
 Brownian motion, 43, 298
 Buckingham, R. A., 151
 Bunsen, R. W., 95
 Bush, W., 315
- Cajori, F., 93
 Calcite, schematic crystal, 49
 Caldwell, S. H., 315
 Callen, H. B., 296
 Capel, W. H., 383
 Carbon, 246
 - tetrahedral structure of, 248
 - trigonal structure of, 248
 Cassels, J. M., 279
 Chadwick, J., 57, 100, 106, 159
 Chain reactions, molecular, 243
 Chapman, S., 342, 387
 Charge, on a condenser, fluctuation of, 297
 - elementary atomic, 51
 Charge-to-mass ratio, electron, 55
 - ions, 55
 - proton, 59
 Charges of atomic components, 50
 Chemical reactions, 243
 Chu, D. Y., 133
 Clapeyron's equation, 18
- Classical electron radius, 31, 390
 Classical foundations of atomic theory, 3 ff.
 Classical physics, 3, 5
 Classical statistics, applications of, 293
 Cleeton, C. E., 113
 Clews, C. J. B., 64
 Coblenz, W. W., 306
 Coefficient, of heat conduction, 340
 - of viscosity, 338
 Coggeshall, N. D., 56
 Cohen, E. R., 50, 51, 65, 84, 117, 133, 390
 Coherent scattering, 152, 225, 228
 - molecular, 231
 Collision broadening of spectrum lines, 109
 Commutator, operator, 130
 Complementarity, 69, 130
 Compounds, organic, 244
 Compressibility, crystal, 359, 360
 Compton, A. H., 62, 93, 159, 279
 Compton effect, 63
 Compton scattering, 153
 Compton wavelength, 63, 134, 390
 Condon, E. U., 142, 159, 206, 210
 Conduction band, 368
 Conductors, 369
 Configuration, electron, 165
 Conservation, of angular momentum, 10
 - relativistic, 173
 - of energy, 11
 - relativistic, 22
 - of momentum, 8
 Continuum of states, 141
 Coolidge, A. S., 239
 Coolidge, J. L., 321
 Coordination number, 350
 Coordination shell, 350
 Coulomb, 23
 Coulomb integral, 216
 Coulomb's law, 23, 24
 Coupling, electronic, 178
 - Russell-Saunders, 163, 165, 178, 208
 Covalent binding, 202
 Covalent bond, 239
 Cowling, T. G., 342, 387
 Crane, H. R., 76
 Crew, H., 93
 Critical angle, X-ray, 47
 Critical pressure, 352
 Critical temperature, 352
 Crookes, W., 52
 Cross, W. G., 153
 Crystal, atomic properties of, 358
 - ideal, 358
 - index of refraction of, 47
 - real, 358

- Crystal lattice energy, 359, 361
 Curie law, 332
 Curie temperature, 375
 Curved grating, 120
 Cyclotron frequency, 58, 74
 Czerny, M., 258
- Dailey, B. P., 334
 Dalton, J., 41
 Damm, C. C., 59
 Darrow, K. K., 93, 357
 Davison, C. M., 65, 151
 de Broglie, L., 67
 Debye, P. J., 229, 362
 Definite and multiple proportions, chemical laws of, 40
 Degeneracy, 212
 DeHaas, W. J., 333
 Democritus, 40
 Density of states, 144
 Detailed balancing, principle of, 308
 Deutsch, M., 154
 Diamagnetic effect, 330
 Diamagnetic susceptibility, 331
 Diatomic molecular constants, 244
 Diatomic molecules, 251
 Dicke, R. H., 113
 Dicke, G. H., 265
 Diffraction, 118
 X-ray, 44
 Diffraction grating, 44, 119
 Diffraction spectrometers, 119
 Dipole energy, electric, 86
 Dipole moment, 145
 electric, 27
 Dipole radiation, classical, 27
 quantum, 145
 selection rules for, 146
 Dirac, P. A. M., 77, 136, 159, 167, 223, 313
 Dirac equation, 168
 Dirac matrices, 171
 Dispersion, 118, 119, 151
 Dispersion spectrometers, 118
 Dissipative systems, 295
 Dissociation energy of hydrogen, 268
 Distance of closest approach, 105
 Doppler effect, 108
 d'Or, L., 326
 Driscoll, R. L., 83
 Dubridge, L. A., 94
 DuMond, J. W. M., 50, 51, 64, 65, 84, 117, 390
 Dunnington, F. G., 55
 Dushman, S., 387
 Dynamics, relativistic, 21
- e/m for electrons, 55
 e/M for positive ions, 55
 e/M for proton, 59
 Eigenfunctions, 131
 energy, 124
 spin, 182
 Einstein, A., 39, 43, 61, 98, 308, 362
 Einstein coefficients, 309
 Electric dipole energy, 86
 Electric dipole moments, 27, 334
 Electric field, 23
 Electric monopole energy, 86
 Electric monopole moment, 87
 Electric potential, 23
 Electric quadrupole moment, 87, 148
 Electric susceptibilities, 329, 333
 Electrodynamics, relativistic formulation of, 34-36
 Electromagnetic momentum, 36
 Electromagnetic oscillators, 28ff.
 Electromagnetic radiation, absorption of, 143, 147
 classical description of, 136
 of energy, 30
 Electromagnetic waves, 26, 28ff.
 Electromagnetism, 23-26
 Electron, 52
 charge of, 390
 classical radius of, 134
 intrinsic spin of, 76, 160
 magnetic moment of, 75
 in a periodic field, 365
 rest mass of, 55, 390
 Electron affinity, 205
 Electron atomic weight, 390
 Electron band, 368
 Electron band theory, 364
 Electron binding energies, 204
 Electron conduction, 368
 Electron configuration, 202
 Electron current, fluctuations in, 299
 Electron cyclotron frequency, 83
 Electron diffraction by crystals, 64, 65
 Electron gas in a metal, 378
 Electron scattering, 107
 Electron shells, 201
 Electronic angular momentum, 241, 252
 Electronic band structure, 370
 Electronic coupling, 178
 Electronic molecular spectra, 236, 266
 Electrostatic energy, of two charges, 124
 two-electron atom, 214
 Electrostatic units, 389
 Ellis, C. D., 100, 159
 Emission of photons by atoms, 143
 Energy, activation, 337
 binding, 202, 242

- Energy, concept of, 10
 conservation of, 11, 22
 crystal lattice, 359, 361
 dissociation, 243
 electric dipole, 86
 electromagnetic radiation of, 30
 electrostatic, 124, 214
 equipartition of, 294
 free, 17
 general conservation equation, 35
 inertial, 11
 internal, 15
 kinetic, 10, 11
 magnetic, 124
 potential, 10, 253
 relativistic transformation of, 23
 resonance, 248
 Energy density, electromagnetic, 29
 Energy eigenfunctions, 124
 Energy eigenvalues, 124
 Energy levels, experimental determination of, 108 ff.
 molecular, 233
 quantitative methods for, 210
 Energy operator, 124
 Energy quantization, 131
 Ensemble, 286
 Enthalpy, 17
 Entropy, 16, 293
 Epicurus, 40
 Epstein, P. S., 39
 Equipartition of energy, 294
 Eriksson, H. A. S., 216
 Estermann, I., 66
 Eucken, A., 326
 Evans, R. D., 151
 Ewald, H., 56, 58, 89
 Ewen, H. L., 113
 Exchange integral, 217
 Eyring, H., 387
- Faraday, M., 23, 24, 50
 Faraday (unit), 51, 390
 Felt, G. L., 64
 Ferguson, A., 296
 Fermi, E., 190
 Fermi-Dirac distribution, 301
 Fermi-Dirac statistics, 301
 Ferroelectricity, 372
 Ferromagnetism, 372
 exchange forces, 376
 Feynman, R. P., 315
 Field emission effect, 381
 Fine structure, of helium, 197
 hydrogenic, 178
 Fine-structure constant, 63, 134
 Fine-structure constant, reciprocal, 390
 Finkelnburg, W., 93
 Fluctuation of charge on a condenser, 297
 Fluctuations in electron current, 299
 Fock, V., 224
 Foley, H. M., 77, 190, 195
 Forbidden band, 368
 Force, concept of, 6
 electric, 23
 electromagnetic, 23
 exchange, 217
 of gravitation, 8
 Form factor, 230
 Fortrat diagram, 270
 Foster, E. W., 188
 Fourier transforms, 15
 Fowler, R. H., 151, 321, 387
 Fowler, W. A., 57, 89
 Franck, J., 99
 Franck-Condon principle, 267
 Frank, N. H., 39
 Fraser, R., 93, 159
 Fraunhofer, J. von, 95
 Fraunhofer lines, 98
 Free radical, 243
 Frenkel, J., 387
 Friedenstein, H., 378
 Froome, K. D., 115
 Fry, I. C., 321
 Furth, R., 296
- Galilean system, 18
 Gardner, J. H., 82, 83
 Garstens, M. A., 80
 Gas, 323
 Gas constant, 390
 Gas laws, 41, 310
 Gas reaction rates, 337
 Gay-Lussac, 41
 Gaydon, A. G., 108, 159, 206, 279
 Geiger, H., 99, 103
 Gerlach, W., 77
 Germer, L. H., 65
 Gibbs, J. W., 286, 321
 Gibbs free energy, 17
 Gibbs-Helmholtz equations, 37
 Gingrich, N. S., 233, 350
 Glassstone, S., 387
 Goldstein, H., 39, 74, 251
 Gordy, W., 113
 Goudsmit, S. A., 59, 75, 159, 162, 190,
 207, 222
 Gram equivalent, 50
 Grating space measurement of crystal, 45
 Graves, A. C., 82
 . Gravitational constant, 390

- Gray, L. H., 151
 Green, A. E. S., 64
 Green, H. S., 350, 352
 Griffith, B. A., 251
 Grotrian diagram, 97, 196
 Guggenheim, E. A., 321
 Gurney, R. W., 321, 360, 387
 Gyromagnetic effect, 373
 Gyromagnetic ratio, 373
- h*, Planck's constant, 61
h/e, 62
 Hagstrum, H. D., 244
 Hamilton, D. R., 110, 194
 Hamiltonian energy function, 12
 Hamiltonian operator, 123
 Hansen, W. W., 81
 Harmonic oscillator, 137
 molecular, 254, 257
 Harnwell, G. P., 39, 94
 Harris, J. N., 64
 Harrison, G. R., 115, 159
 Hartley, W. N., 95
 Hartree, D. R., 223, 224
 Hays, E. E., 59
 Heat, 15
 latent, 244
 mechanical equivalent of, 390
 molecular (*see* Molecular heat)
 specific (*see* Specific heats)
 Heat conduction, coefficients of, 342
 Heisenberg, W., 70, 122
 Heitler, W., 136, 150–152, 159
 Helium, ground state of, first approximation, 216
 liquid, 355
 Helium-atom problem, 212
 Helium-energy-level scheme, 217
 Helium II, 355
 Helmholtz free energy, 17
 Hermite polynomials, 138
 Herring, C., 378
 Hertz, G., 99
 Herzberg, G., 159, 206, 244, 260, 279
 Herzog, G., 229
 Heteropolar binding, 240
 Hewett, W. H., 373
 Hippel, J. A., 58, 83, 84
 Hofstadter, R., 153
 Homonuclear molecules, 274
 Homopolar binding, 239
 Honig, H., 259
 Hooke's law, 13
 Hopper, V. P., 52
 Houston, W. V., 39, 133, 159, 210, 223
 Hughes, A. L., 94, 387
- Hulme, H. R., 151
 Hund's rule, 376
 Huxford, W. S., 297, 299
 Hydrogen atom, polarizability of, 226
 relativistic, 173
 HCl absorption bands, 258–260
 Hydrogen iodide, reaction rate, 338
 Hydrogen molecule, 236
 configurations, 238
 Hydrogen spectral series, 95
 Hydrogen wave functions, 132
 Hylleraas, E. A., 216
 Hyperfine structure, 188
 Hyperfine-structure separations, 195
- Ice point, 390
 Identical particles, 197
 Impact parameter, 104
 Imperfect gas parameters, 349
 Imperfect gases, 342
 Incoherent scattering, 225
 Index of refraction of a crystal, 47
 Induced emission, 145, 308
 Inertial energy, 11
 Inertial reaction, 11
 Infeld, L., 39
 Infrared spectra, 257
 Inherent line breadth, 110
 Insulators, 369
 Intensity of radiation, total, 145
 Internal energy, 17
 Intrinsic angular momentum, 75
 Ionic binding, 240
 Ionic crystals, 358
 Ionic radii, 360
 Ionization potential, 204, 205
 Ionization process, 98
 Ions, 52
 Isentropic process, 17
 Isoelectronic sequences, 108
 Isothermals, 352
 Isotopes, 53, 55
 Isotopic weight, 49
 Isovoluminal isentropic process, 17
 Ives, H. E., 20
- Jackson, D. H., 188
 James, H. M., 239
 Jeffries, C. D., 84
 Jenkins, R. O., 378
 Jevons, W., 279
jj coupling, 210
 Johnson, F. T., 64
 Johnson, T. H., 66
 Jones, H., 387

- Joos, G., 39, 70
 Jordan, E. B., 56
- Karplus, R., 78
 Keesom, W. H., 352
 Keesom's equation, 275
 Kellogg, J. B. M., 110, 194
 Kennard, E. H., 94, 321, 387
 Kennedy, R. J., 19
 Kent, N. A., 188
 Keynes, J. M., 321
 Kikuchi, S., 66
 Kilogram, 7
 Kilomole, 42
 volume of one, 390
 Kinematics, 6
 relativistic, 21
 Kinetic energy, 10, 11
 Kirchhoff, G. R., 95
 Kistiakowski, G., 338
 Kittel, C., 372, 373
 Klein, O., 153
 Klein-Nishina formula, 153
 Knight, W. D., 80
 Koenig, S., 77
 Kroll, N., 78
 Krutter, H. M., 315
 Kuhn, H., 188
 Kusch, P. 77, 82, 83, 190, 194, 195
 Kyhl, R. L., 113
- Laby, T. H., 52
 Lagrangian function, 12
 Lagrangian undetermined multipliers, 288
 Laguerre's equation, 131
 Laidler, K. J., 387
 Lamb, W. E., Jr., 112, 179, 195, 197, 320
 Lamb shift, 195, 196, 221
 Landé interval rule, 179, 190
 Landé splitting factor, 163
 Landsberg, P. T., 377
 Laplace distribution law, 284
 Larmor frequency, 74, 79
 Latent heat, 244
 Lauritsen, C. C., 57, 89
 Law of atmospheres, 295
 Le Corbeiller, P., 222
 Leucippus, 40
 Leventhal, E. C., 82
 Li, C. W., 57, 89
 Light, velocity of, 19, 115, 390
 Lindsay, R. B., 321
 Line breadth, natural, 149
 Lipworth, E., 197
- Liquid helium, 355
 Liquid state, 349
 Location of electrons and nuclei in molecule, 224
 London, F., 355, 387
 Loofbourow, J. R., 115, 159
 Lord, R. C., 115, 159
 Lorentz transformation, 20
 Lorentz triplet, 165
 Louisell, W. H., 76
 Lutze, E., 151
 Lyman, T., 96
- McDouglass, J., 151
 McIntyre, J. A., 153
 Mack, J. E., 188, 195
 McMillen, J. H., 233
 Madelung, E., 359
 Magie, W. H., 94
 Magnetic deflection of photoelectrons, 64
 Magnetic dipole energy, 86
 Magnetic dipole moment, 87
 Magnetic effect, simple, 160
 Magnetic energy, 124
 Magnetic induction, 23
 Magnetic moment-angular momentum ratio, 71, 163
 (See also Gyromagnetic ratio)
 Magnetic moments, 24, 70, 72
 electron spin, 78
 nuclear, 79, 82, 89, 190
 Magnetic quantum number, 132
 Magnetic-resonance absorption, 80
 Magnetic susceptibilities, 329
 Magnetic vector potential, 25
 Magnetism, 370
 March, A., 159
 Margenau, H., 169, 347
 Marney, M. C., 66
 Marsden, E., 100
 Martin, S. L., 378
 Mass, of atomic components, 50
 concept of, 7
 gravitational, 8
 inertial, 8
 neutron, 57
 from nuclear transmutations, 56
 relativistic, 21
 Mass-doublet measurements, 56, 57
 Mass spectrographic analysis of ionization products, 108
 Mass-spectrographic determination of masses, 57
 Mass spectroscopy, 55
 Mass synchrometer, 59
 Massey, H. S. W., 279

- Matrices, 168
 alpha, 171
 conjugate, 172
 Matrix elements, 170
 Matrix order, 172
 Matter, wave nature of, 65
 Maxwell's equations, of electricity and electromagnetism, 24
 of thermodynamics, 36
 Maxwell-Boltzmann velocity distribution, 336
 Mayer, J., 387
 Mayer, M. G., 387
 Mayne, K. I., 56
 Mean free path, 336
 Mechanical equivalent of heat, 390
 Mechanical properties of crystals, 357
 Mechanics, Newtonian, 6 ff.
 Mees, C. E. K., 39
 Meissner, K. W., 189
 Mendeleev, D. I., 41
 Mendelssohn, K., 357
 Merritt, F. R., 373
 Metals, 358
 Metastable state, 99
 Methane, 245
 Metropolis, N., 315
 Microwave absorption, 114
 Microwave interferometer, 116
 Microwave spectrometer, 113
 Microwave technique, 112
 Millikan, R. A., 52, 61, 94
 Millman, S., 110, 194
 mks units, 6, 389
 Mole, 42
 Molecular angular momentum, 252
 Molecular-beam technique, 79, 110, 111
 Molecular binding, 239
 Molecular binding energy, 242
 Molecular bonds, 249
 Molecular constants, diatomic, 244
 Molecular crystals, 358
 Molecular diatomic potential-energy curve, 242
 Molecular electronic states, 233, 236, 240
 Molecular heat, electronic, 326
 hydrogen, 329
 rotational, 327
 translational, 325
 vibrational, 326
 Molecular nuclear motion, 249
 Molecular potential energy, 253
 Molecular rotation frequencies, 235
 Molecular rotation spectra, 258
 Molecular spectra, 236, 266
 nuclear spin in, 271
 Molecular spectroscopy, 255
 Molecular structure, 224 ff.
 Molecular symmetry, 240
 Molecular vibration frequencies, 235
 Moments, antiferromagnetic, 372
 high-order, 85
 multipole, 87
 quadrupole, 87, 148
 Momentum, concept of, 7
 conservation of, 8
 generalized, 36
 relativistic transformation of, 23
 (*See also* Angular momentum)
 Moore, C. E., 179, 206, 223
 Morse, P. R., 142, 159, 250, 357
 Morse molecular potential function, 242,
 250
 Moseley, H. G. J., 101, 106
 Mott, N. F., 159, 223, 279, 360, 387
 Mulligan, J. F., 115
 Multiplet separations, 179
 Multiplicity, spin, 166
 Mundan, G. L., 378
 Mundie, L. G., 189
 Murphy, G. M., 169
 Nafe, J. E., 77
 Natural line breadth, 148
 Nelson, E. B., 77
 Neutron, 56
 Neutron mass, 57
 Neutron scattering, 373
 Newton, I., 8, 95
 Nichols, M. H., 378
 Nicodemus, D., 82
 Nier, A. O., 56
 Nishina, Y., 153
 Noise voltage, 297
 Normal distribution, 284
 Normal-error curve, 284
 Normal Lorentz triplet, 162
 Normal modes of radiation, 305
 Normalizing condition, 122
 Novick, R., 82, 197
 Nuclear atom, 99
 Nuclear magnetic moments, 79, 82, 89
 Nuclear magnetic resonance, 80
 Nuclear magneton, 82
 Nuclear quadrupole moments, 89
 Nuclear resonance induction, 81
 Nuclear spin in molecular spectra, 271
 Nucleons, 79
 Nuttall, J. M., 151
 Nyquist, H., 296
 Omegatron, 58
 One-electron spectra, 161

- Onnes, K., 332
 Opalescence near critical point, 352
 Oppenheimer, J. R., 233
 Optical spectrum region, 117
 Orbital overlapping, 246
 Orbital quantum number, 132
 Organic compounds, 244
 Ornstein, L. S., 296
 Ortho-hydrogen, 328
 Orthonormality, 139, 140, 210
 Oscillator strength, 226
 Oscillators, harmonic, 137
 molecular, 254, 257
- Packard, M. E., 81, 82
 Page, L., 39
 Pair production, 154
 Pake, G. E., 80, 194
 Panofsky, W. K. H., 64
 Para-hydrogen, 328
 Paramagnetic susceptibilities, crystal, 372
 Paramagnetic susceptibility, 331
 Paramagnetism, 330
 crystal, 371
 Parity, 128
 Parkinson, W. C., 153
 Partition function, 289, 343
 intramolecular, 344
 Paschen, F., 96, 217
 Paschen-Back effect, 165, 187
 Pauli principle, 197-201
 Pauli spin matrix, 171
 Pauling, L., 138, 159, 223, 237, 279
 Pearse, R. W. B., 279
 Pellam, J. R., 357
 Perfect differential, 16
 Perfect gases, 16, 312, 324
 Periodic table, 42, 200
 Permittivity of space, 390
 Perrin, J., 42, 94
 Perturbation method, 140
 Perturbation theory, 211
 time-dependent, 138
 Perturbed molecular levels, 269
 Pfund, A. H., 96
 Phase space, 300
 Phase-space cells, 310
 Phonons, 363
 Photo detectors, 118
 Photoelectric cross section, 151
 Photoelectric effects, 62, 150, 377, 381
 Photons, 61, 85
 Physical measurement, philosophy of, 280
 Physical scale, atomic, 49
 Physical science, 1 ff.
- Pidd, R. W., 76
 Pirenne, M. H., 279
 Planck, M., 61
 Planck's black-body radiation law, 306
 Planck's constant, 61, 390
 Poisson distribution law, 317
 Polarizability, atomic, 225
 hydrogen atom, 226
 Polyatomic molecules, 245
 Polyelectronic atoms, 160, 313
 Polyelectronic wave function, 199
 Positive rays, 55
 Positronium, 154
 Potassium chloride, energy of formation, 361
 Potential energy, 10
 Pound, R. V., 80, 82, 175
 Poynting's vector, 29
 Precession of rotating electromechanical system, 74
 Price, W. C., 118
 Pringsheim, P., 223
 Probability, 69, 281
 Probable error, 284
 Prodell, A. G., 77, 195
 Progressions, spectral, 268
 Propagation vector, 137
 Proper time, 21
 Properties of matter, classical, 322
 elementary, 322
 Proton atomic weight, 390
 Proton-electron mass ratio, 390
 Proton specific mass, 58
 Prout, J. L., 40
 Purcell, E. M., 80, 82, 83, 113
- Quadrupole absorption, 148
 Quadrupole energy, electric, 86
 Quadrupole moment, electric, 87, 148
 Quantitative methods for energy levels, 210
 Quantization, angular-momentum, 126
 energy, 131
 of the radiation field, 135
 Quantum numbers, 132, 138
 Quantum radiation processes, 143
 Quantum statistics, 299
 applications of, 304
 Quantum theory of radiation, 135
 Quantum thermodynamics, 303
- Rabi, I. I., 79, 110, 111, 192, 194
 Radial wave functions, 134, 135
 Radiation, black-body, 304
 from a circulating charge, 32

- Radiation, corpuscular nature of, 60
 electromagnetic (*see* Electromagnetic radiation)
 modes of, normal, 305
 quantum or photon hypothesis of, 61
 quantum theory of, 135
 total intensity of, 145
- Radiation damping, 31
- Radiation processes, 143
- Radical, free, 243
- Raman, C. V., 262
- Raman effect, 151
- Raman scattered components from HCl, 265
- Raman scattering, 152, 263
- Raman selection rules, 263
- Raman spectra, 262
- Ramsey, N. F., 89, 94, 112, 153, 190, 194
- Ramsey's arrangement, 112
- Randall, J. T., 279
- Randomness, 281
- Rate processes, 335
- Reactions, chain, 243
- Real gas, 348
- Reciprocal fine-structure constant, 390
- Relativistic formulation of wave equation, 166
- Relativistic mechanics, 18 ff.
- Residual ray, 362
- Resolving power, 118
- Resonance energy, 248
- Resonance fluorescence, 152
- Resonance methods for determining e/m , 58
- Resonance scattering, 263
- Reststrahlen, frequencies of, 362
- Retarded potentials, 26
- Rutherford, R. C., 195
- Rice, O. K., 279
- Rice, S. O., 296
- Richards, R. I., 59
- Richardson-Dushman equation, 381
- Richter, J. B., 40
- Richtmyer, F. K., 94
- Riemann, A. L., 387
- Ritz combination principle, 96
- Roberts, A., 82
- Roberts, T. R., 56
- Robertson, H. P., 39
- Robinson, C. F., 133
- Robinson, H. R., 64
- Roentgen, K., 52
- Rogers, E. H., 82
- Rollin, R. V., 80, 188
- Rotational levels, molecular, 256
- Rowland, H. A., 95
- Rowland circle, 119
- Rubinowicz, A., 223
- Rupp, E., 66
- Russell-Saunders coupling, 163, 165, 178, 208
- Rutherford, E., 99, 100, 159
- Rutherford scattering, 103
 cross section of, 105
- Rydberg constant, 96, 133, 390
- Sadler, C. A., 101
- Saha's equation, 386
- Sawyer, R. A., 115, 159
- Scattering, α -particle, 100
 coherent, 152, 225, 228, 231
 Compton, 153
 Coulomb, 101
 of electromagnetic wave, 32
 electron, 107
 incoherent, 225
 of light, 352
 multiple, 103
 neutron, 373
 normal modes of, 305
 Raman, 152, 263
 resonance, 263
 Rutherford, 103, 105
 Thomson, 34, 102
 X-ray, from liquid argon, 350
- Scattering cross section, classical, 33
- Scattering factor, 232
- Schiff, L. I., 142, 159, 185, 210, 223
- Schottky effect, 381
- Schroedinger, E., 121, 321, 387
- Schroedinger theory, 121
- Schroedinger's equation, 121
 periodic perturbation, 365
- Schwinger, J., 77
- Scientific theory, 1 ff.
- Second (unit of time), 6
- Seitz, F., 358, 387
- Selection rules, for dipole radiation, 146
 for electronic levels, 256
 in magnetic field, 160
 for molecular levels, 274
 Raman, 263
 for vibrational levels, 256
- Semiconductors, 358
- Separation of variables, 125
- Series, spectral, 97, 206
- Shannon, C. E., 298
- Shenstone, A. G., 108, 205
- Shields, R. B., 56
- Short-wavelength limit, X-ray, 64
- Shortley, G. H., 159, 206, 210
- Shot effect, 299
- Showers, cosmic rays, 154

- Shull, C. G., 373
 Shull, E. G., 66
 Shulman, R. G., 334
 Silverman, S., 296
 Simple rotator, molecular approximation, 251
 Singlet term, 213
 Size of an atom, 42
 Skinner, H. W. B., 377
 Slater, J. C., 39, 204, 206, 224, 315, 387
 Smabler, B., 82
 Smart, J. S., 373
 Smekal, A., 262
 Smith, L. G., 59
 Smyth, H. D., 108
 Sneddon, I. N., 159, 223, 279
 Sommer, H., 58, 84
 Spatial localization, 68
 Specific electronic charge, 390
 Specific heats, 324
 curves of, 364
 of a crystal, 361
 electron contribution to, 379
 Spectral series, 97, 206
 Spectrum-line widths, 108
 ultimate, 110
 Spectrum patterns, X-ray, 107
 Spence, R. W., 82
 Spherical electromagnetic wave, 28
 Spin, 76, 160, 162, 175
 Spin eigenfunctions, 182
 Spin-orbit interaction, relativistic, 175, 177
 Spontaneous emission, 145
 Sproull, W. T., 279
 Standard deviation, 283
 Standards of length, 6
 Statistics, of atomic processes, 280
 of observation, 280
 Staub, H. H., 82
 Stefan's law, 320
 Stefan-Boltzmann law, 307
 Stern, O., 66, 77, 382
 Stern-Gerlach experiment, 77
 Stillwell, G. R., 20
 Stirling's approximation, 283
 Stitch, M. L., 259
 Stockmayer, W. H., 348
 Stokes lines, 265
 Stokes's law, 265
 Stoner, E. C., 372
 Stoney, J., 52
 Stopping potential, 62
 Structure factor, 228
 Susceptibility, magnetic, 371
 Symmetric wave functions, 199
 Synge, J. L., 251
 Taub, A. H., 82, 83
 Teller, E., 315
 Temperature, 15
 statistical significance of, 292
 Term differences, 96
 Term series, 98
 Term values, alkali-metal, 207
 Terms, from electron configurations, 209
 spectroscopic, 96, 165
 Tesseral harmonic, 126
 Thermal density fluctuations, 354
 Thermal equilibrium, 285, 286
 Thermal properties of crystals, 357
 Thermionic effects, 377
 Thermionic emission, 381
 Thermodynamic functions, 17, 293
 Thermodynamics, 15 ff.
 first law of, 15
 quantum, 303
 relation to statistics, 292
 second law of, 15
 Thode, H. G., 56
 Thomas, H. A., 58, 83, 84
 Thomas-Fermi atom, 313
 Thomas-Fermi function, 316
 Thomas-Fermi radial charge density, 316
 Thomson, G. P., 52, 66, 94
 Thomson, J. J., 52, 94, 101
 Thomson atom, 101
 Thomson scattering, 34, 102
 Thorndyke, E. M., 19
 Ting, Y., 113
 Tolman, R. C., 321
 Torrey, H. C., 80
 Total quantum number, 132
 Townes, C. H., 259, 334
 Trajectories of charged particles, 53
 Transfer processes, 335
 Transformation equation, relativistic electromagnetic, 35
 Transition probability, 141, 142
 Transmission line, 296
 Transmutation cycles for nuclear mass differences, 57
 Triplet term, 213
 Two-quantum process, 263
 Uhlenbeck, G. E., 75, 162, 296
 Ultraviolet catastrophe, 61
 Uncertainty, conditions of, 69
 Uncertainty principle, 69
 Units, atomic physics, 389
 definition of, 6
 Unsaturated double bond, 249
 Unsold, A., 149

- Valence, 50
Valence crystals, 358
van der Waals' equation, 348
van der Waals forces, 346
Van Vleck, J. H., 80, 94, 333, 372, 387
Vane, A. B., 113
Velocity, 6
 of light, 19, 115, 390
Velocity-of-light apparatus, 117
Verster, N. F., 94
Vibrational levels, molecular, 256
Virial coefficient, second, 348
Viscosity, 338
Viscosity coefficients, 342
Volume of one kilomole, 390
von Friesen, S., 66
- Walcher, W., 56
Wang, M. C., 296
Water-forming reaction, 243
Water molecule, 246
Watts, H. M., 64
Wave equation, relativistic, 168
Wave functions, angular, 127
 antisymmetric, 199
 hydrogen, 132
 localization of, 67
 polyelectronic, 199
 radial, 134, 135
 symmetric, 199
Wave motion, 12 ff.
Wave nature of matter, 65
Wave number, 96
Wave properties of material particles, 66
Weiskopf, V. F., 39, 152
Weiss, P. R., 374
Welton, T. A., 221, 296
Whaling, W., 57, 89
White, H. E., 108, 159, 206, 223
Whitehead, A. N., 39
Wien's displacement law, 307
- Wiersung, E. C., 333
Williams, D., 113
Williams, E. J., 151
Williams, N. H., 113, 297, 299
Williams, R. C., 133
Williams, Van Z., 118
Wilson, E. B., 138, 159, 223, 237, 279
Wollan, E. D., 66, 229, 233
Wollaston, W. H., 95
Woltjer, H. R., 332
Wood, R. W., 265
Work, 10
Work functions, metallic, 382
Worsnop, B. L., 66
- X-ray diffraction, 44
X-ray line structure, 377
X-ray region, 118
X-ray scattering from liquid argon, 350
X-ray spectra, 106
- Yager, W. A., 373
Yasaitis, E., 82
Yergin, P. F., 197
Young, L. B., 56
- Zacharias, J. R., 194
Zachariasen, W. H., 94
Zahn, C. T., 384
Zeeman effect, 161
 initial, 163, 187
 one-electron, 180
Zeeman multiplet polarization, 165
Zeeman splitting, 185
Zemansky, M. W., 39
Zeno's paradox, 40
Zero-moments method, 193, 194
Zero-point oscillation, 197, 256
Zinn, W. H., 66