INTERNATIONAL SERIES IN PURE AND APPLIED PHYSICS
G. P. HARNWELL, CONSULTING EDITOR

ATOMIC PHYSICS

INTERNATIONAL SERIES IN PURE AND APPLIED PHYSICS

G. P. HARNWELL, CONSULTING EDITOR

Allis and Herlin Thermodynamics and Statistical Mechanics
Cady Piezoelectricity
Clark Applied X-rays
Edwards Analytic and Vector Mechanics
Finkelnburg Atomic 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
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
Richtmyer and Kennard Introduction to Modern Physics
Ruark and Urey Atoms, Molecules, and Quanta
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
Stratton Electromagnetic Theory
Thorndike Mesons: A Summary of Experimental Facts
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.

Dr. Lee A. DuBridge was Consulting Editor of the series from 1939 to 1946.

ATOMIC PHYSICS

By

WOLFGANG FINKELNBURG, Ph.D.

Formerly Professor of Physics, University of Strassburg

Translated from the Revised German Edition
in Cooperation with the Author
by George E. Brown, M.Sc.

NEW YORK TORONTO LONDON
McGRAW-HILL BOOK COMPANY, INC.
1950

ATOMIC PHYSICS

II

PREFACE

Atomic physics, as treated in this book, deals with our entire knowledge of the structure of matter, from the elementary particles up to the solid state. It is presented here for seniors and graduate students of physics, chemistry, and engineering, and for those professional workers who want to keep pace with the rapid progress in this important field. It is hoped that the book will lead the reader to a true understanding of this vast field which has changed our concepts of nature so fundamentally during the last few decades and has found such an amazing number and variety of applications. This book therefore stresses the meaning of experiments and theories, and the interrelations between such seemingly divergent fields as atomic, molecular, nuclear, and solid-state physics rather than mathematical or experimental details. These are treated only to that extent which was believed necessary for a thorough understanding. However, it is intended to lead the reader right up to the present front line of atomic research and enable him to appreciate the far-reaching problems and implications which are at stake.

The book is a translation from the German "Einführung in die Atomphysik," published first in 1948 by Springer-Verlag, and is essentially identical with the presentation in the revised second German edition which is now in preparation. Every possible effort has been made to bring the book up to date, and all available published material up to the spring of 1949 has been taken into consideration.

The author is greatly indebted to a large number of colleagues and their agencies for having so generously supplied excellent photographs and diagrams used in this book. Acknowledgments are made in the captions of the figures. The author also wishes to express his deep appreciation and gratitude for the excellent collaboration of his translator, George E. Brown.

WOLFGANG FINKELNBURG

ENGINEER RESEARCH AND
DEVELOPMENT LABORATORIES
FORT BELVOIR, VA.
March, 1950

CONTENTS

CHAPTER 1

INTRODUCTION

In this book atomic physics is presented in a broad sense. It deals with our knowledge of the structure of matter, its constituent elementary particles, its properties and behavior. Atomic physics attempts to explain the immense number of physical phenomena of our world by means of a few general laws and fundamental constants of nature and from the properties of a very few elementary particles. This science of the "atomistic" structure of atomic nuclei, atoms, molecules, and crystals, though it has its roots in the last century, is, actually, the result of twentieth-century physics. The beginning of this century is highlighted by Planck's discovery of the elementary quantum of action h and the realization of its universal importance. This led to the development of the quantum theory, which provided the key to the understanding of all phenomena of atomic physics. A large percentage of all physical research of our century is in some way concerned with problems of atomic physics.

1-1. The Significance of Atomic Physics in Science and Its Practical Applications

The new concept of matter, and hence of the foundations of physics and chemistry, has greatly altered our concepts of many older and well-known fields of physics. For this reason we speak of modern physics in contrast to classical physics. Classical physics in this sense means our knowledge of physics as scientists of the last century had perfected it without taking into account any atomic phenomena. The treatment of physics with the aim of explaining *all* physical phenomena, whether mechanical, thermal, electrical, or optical, from the fundamental atomistic point of view, is called "modern physics."

Let us see how fundamentally our outlook on whole fields of physics has been changed and clarified by regarding physical phenomena from an atomistic point of view. For instance, in the last century a beautiful and formally complete thermodynamic theory of heat had been developed. Nevertheless, the basic thermal processes were not understood until the thermodynamic theory was supplemented by an essentially atomistic theory, the kinetic theory of heat. In a similar way modern atomic physics has explained metallic conductivity, thus solving one of the fundamental problems of electricity. The phenomena of elasticity, thermal expansion, and many other features of solid matter, phenomena near the absolute zero of temperature, and many other problems of macrophysics can be understood much more satisfactorily from the basis of atomic physics. The same is true for the broad field of optics. In addition to contributing to a fundamental understanding of many of the long-known fields of macrophysics, atomic physics has also opened entirely new fields in science. These fields, which are concerned with the structure of atomic nuclei, atoms, molecules, and crystals, are discussed in this book. It is truly a tremendous unification and expansion of our physical picture of the world which has been produced by atomic physics.

However, the importance of atomic physics is not confined to physics proper. The development of many other fields of science and industry, in fact, our whole concept of the world, has been deeply influenced by it. The closely allied science of chemistry has received fundamental contributions in the form of the explanation of the periodic system of elements and the theory of chemical binding. Moreover, it uses to an ever increasing extent new methods of research developed in molecular physics. Major parts of astrophysics today are actually applied atomic physics, in particular, applied spectroscopy. Spectroscopic methods enable the astronomer to study stellar atmospheres, temperatures, and distances. Moreover, the problem of stellar energy production was solved by the discovery of suitable exothermal nuclear reactions, and this explanation opened the road to a better understanding of stellar evolution and cosmology. Crystallographers and mineralogists utilize to a great extent experimental as well as theoretical methods of atomic physics. By means of optical spectra and X-ray diffraction they study the constitution and structure of crystals and minerals, and they use theoretical results on atomic properties and binding forces in their attempts to understand mixed crystal formation and similar phenomena. Biologists study their organisms with high-powered electron microscopes, and irradiate them with ultraviolet light, X-rays, or neutrons to investigate their reactions and mutations. They use radioactive isotopes for the study of metabolism and have received new stimulation from Jordan's quantum-biological ideas, the future importance of which is still difficult to estimate. Furthermore, there seems to be little doubt that in a future physicochemical explanation of the autocatalytic processes of cell proliferation (chromosome duplication) the results of molecular physics will play a decisive role. Even mathematicians have been stimulated by the latest developments of quantum mechanics to discuss more thoroughly the possibilities of a mathematical formalism using finite smallest quantities, which would be better adapted to the discontinuous quantum phenomena than differential equations using infinitely small changes of all quantities.

Also of considerable importance is the influence of atomic physics on many fields of industry of which we want to cite only a few examples. Modern technology of gas discharges and light sources is almost completely applied atomic physics. The same is true for modern communication techniques and electronics, since the construction of the numerous electronic instruments from oscilloscopes, image tubes, television equipment, thyratrons, and radio equipment up to the electron microscopes is based on pure atomic research. Many fields of electrical engineering (such as rectifier technique or electrical contacts) still require atomic research for their further development, and the same applies to the improvement of ferromagnetic and other materials. The properties of all solid materials depend on interatomic forces, so that all progress made in atomic physics of the solid state will have its future effect on this important field of technology. Nondestructive testing of materials by spectroscopic or X-ray methods plays an increasingly important role in this same field. The wide use of phototubes and photosensitive elements in industry and the practical importance of molecular radiation for determining flame-gas behavior in modern combustion engines indicate the variety of applications of atomic physics in modern technology. There is no doubt whatsoever that the extent of the application of atomic physics will increase greatly, so that a sound understanding of atomic physics will be a requirement for the future engineer.

Something should be said, in concluding this discussion, of the implications of atomic physics with respect to our general concepts of the natural sciences and even philosophy itself. Its most important effect is that it has fundamentally changed our all too rigidly mechanical concepts of the physical quantities, matter and energy; of the meaning of forces; and of the meaning and importance of the philosophical categories, substance and causality. We can notice already the deep influence of this change on our whole physical picture of the world, on natural philosophy based on it, and on the philosophy of science, though opinions are still very much divided. The interested reader is referred to books of Bavink, Frank, Heisenberg, Jordan, Margenau, Planck, and von Weizsäcker. The basic difficulties of these philosophic consequences of quantum physics, which will be discussed later, are, no doubt, considerable, especially for the experimental scientist. Nevertheless, it is necessary to study these problems seriously if we really strive to be natural *scientists*, just as, on the other hand, anyone who wants to discuss modern philosophy intelligently must necessarily study atomic physics seriously.

1-2. Methods of Research in Atomic Physics

Atomic physics, since it deals with the structure of matter, should be the foundation of all macrophysics, which takes matter and its properties simply for granted. However, we know that atomic physics actually was conceived after classical macrophysics had been nearly completely developed. Because of this fact, the methods of research, reasoning, and verification are somewhat different from those employed in the different fields of macrophysics. To a great extent experimental and theoretical results from nearly all domains of macrophysics are used in atomic physics. But the "classical picture" of atomic processes, formed in this way, is then developed further by constant cooperation and supplementation of theory and experiment to achieve a better and better agreement. The final "picture" of the atomic phenomenon usually does not have much similarity with the initial classical model. We always have to keep in mind that, in contrast to macrophysics, the objects of our study (electrons, atoms, ions, or molecules) cannot be observed *directly*. We have to deduce our "picture" of them from their influence on our different instruments of observation and measurement.

The general procedure in atomic research is the following: From the first results of accidental or deliberate observations we conceive of a crude pictorial model of the particle (or process). This model tells us what its characteristics (properties) must be in order that the observed effects follow from it as logical consequences. From these properties of the first model we usually are able to make further predictions as to the behavior of the phenomenon under different experimental conditions. Such experiments are then performed, if possible, and their results lead to a confirmation, change, or refinement of the first model. As soon as a certain agreement on the most essential properties of the model under consideration has been reached, it is studied theoretically. It is important to realize that this first theory is concerned with the *model* and not with the phenomenon itself, though both should agree in their most essential properties. This first quantitative theory allows us to draw quantitative conclusions which can be tested by experiments. The results of the experiments help us to decide in what respects further refinements of the model are required. Thus, step by step, by mutually supplementing each other, experiment and theory lead to an increasingly better approximation to what we might call the physical reality. Each new or improved theory stimulates new experiments, the conception of which, by the way, often requires considerable mental ability. Each exact experiment, on the other hand, which does not result in equally exact agreement with the theoretical prediction instigates a further improvement of the theory.

Moreover, we shall see that in several critical phases during the development of atomic physics, new experiments forced the theoreticians to realize that a very radical and fundamental change of existing theories was necessary. Physicists thus were forced to realize that many laws of classical macrophysics whose applicability to atomic phenomena had been considered self-evident lost their validity and applicability in atomic dimensions. The interaction of theory and experiment thus resulted in the development of fundamentally new theories that considerably changed all our physical reasoning. We shall discuss later a number of impressive examples of this close interaction between experiment and theory which is so characteristic of atomic physics. It implies also that the atomic physicist should not be exclusively an experimenter or a theoretician. An experimenter has to understand at least enough theory to be able to draw conclusions from it and explain his experimental results correctly, whereas a theoretician has to know enough of experiments and experimental possibilities to participate in the discussion of experimental results or to judge for himself the possibilities of testing his theories experimentally.

These particular methods of atomic research imply methods of verification which are somewhat different from those in macrophysics. The correctness of an assertion in the field of atomic physics cannot be proved as directly and unambiguously as in macrophysics because of the unobservability of atomic objects. Moreover, the correctness of an assertion such as that of the existence and universal importance of Planck's quantum of action h is based on the fact that this smallest quantity of action, h, appears in *all* decisive formulas of the whole field of atomic physics. All the experiments of mechanical, optical, photoelectrical, or X-ray-spectroscopic nature, though differing widely in the techniques used, result in the *same* value of this smallest quantum of action, h. In a similar way the "correctness" of an atomic theory like quantum mechanics cannot be proved *directly*. Rather its proof follows from the fact that in the range of its asserted applicability (i.e., in all atomic physics with the possible exception of certain extreme domains of nuclear physics) all experimental results can be accounted for quantitatively, while all verifiable new conclusions from this theory have been confirmed by experiments. This complete consistency and agreement is exactly what we mean when we speak of the "correctness" of a theory.

1-3. Difficulties and Methods of Presentation of Atomic Physics

The particular relation of atomic physics to the fields of macrophysics, which we indicated briefly, implies certain difficulties which the beginner encounters in his attempt to study this field and which will be apparent in this book. We have already mentioned that atomic physics makes continual use of the results of nearly all fields of classical physics. A thorough knowledge of classical physics, therefore, should be the basis for studying atomic physics. The kinetic theory, the phenomena of electrolysis, and the basic facts of chemistry form the foundation on which atomic physics has been constructed. The concept of electron orbits in the atom was derived from the planetary orbits. A knowledge of gyro mechanics was necessary in order to recognize the atom as a system of coupled microscopic gyroscopes. The orbital electrons form electric convection currents and produce magnetic fields. An understanding of the effects of these fields is impossible without the basic knowledge of electromagnetism, whereas an understanding of the emission of radiation by atoms requires knowledge of electromagnetic waves and optics. These few examples show how experimental facts and theoretical concepts from all fields of macrophysics are used in atomic physics, and that theory and experiment are more closely interwoven than in most fields of macrophysics. Undoubtedly this implies difficulties for the beginner. This forced versatility, on the other hand, makes atomic physics so attractive and interesting that it justly has been called the outstanding field of all physics.

It is evident from our introductory discussion that atomic physics is not a well-defined field of science. We can either confine our interest to the basic laws of quantum physics and the structure of atoms, or we can go to the other limit and try to derive nearly all phenomena of physics from the properties of elementary particles and thus include all modern physics in atomic physics. Because of the ambiguity in the definition of atomic physics, courses and textbooks vary more widely in their subject matter and manner of presentation than in most other fields of physics. Some authors prefer, for this very reason, to treat the single subjects of atomic physics rather than the whole field. Such presentations, excellent and complete as they may be in themselves, fail to reveal the internal relations and cross connections between the single domains, so that the student is not easily able to realize the wonderful harmony which, in spite of so many unsolved problems, today unifies physics of nuclei, atoms, molecules, and crystals.

It is the aim of this book to point out as clearly as possible this unity and harmony of all domains of atomic physics. It is assumed that the student has a knowledge of undergraduate physics including the essential facts of the kinetic theory of gases and statistics. A further prerequisite is some basic knowledge of theoretical physics, unless the reader is willing to *believe* the correctness of the necessary derivations and theoretical explanations. However, mathematical derivations have been limited to the extent where any further omissions would leave gaps in the discussion. Every effort has been made to lead the reader to a thorough understanding of the basic concepts and the true meaning of all theories.

The order of presentation of atomic physics in this book follows essentially from didactic considerations. However, it actually follows largely the historic development. Chapter 2 presents, besides the evidence for the atomistic nature of matter and electricity and for the existence of electrons, atoms, ions, nuclei, and photons, those facts about these particles (including isotopy) which are prerequisite for their further treatment. In Chap. 3, the basic relations between atomic structure and atomic spectra (which also hold for molecules discussed in Chap. 6) are dealt with on the basis of the Bohr theory. It is in agreement with the historical development that quantum mechanics (treated in Chap. 4) is not yet used in this chapter. Moreover, this arrangement makes it possible to *prove* to the reader the necessity of introducing wave and quantum mechanics, and makes it evident that every new theory does not "overthrow" the old one, but expands, refines, and finally includes it as a special case. In agreement with historical development and our didactic intention, nuclear physics, which in a systematic treatment should form the basis for atomic physics proper, is discussed in Chap. 5, *after* the treatment of quantum mechanics. A real understanding of nuclear processes presumes some quantum-mechanical knowledge, wheareas quantum mechanics itself is shown to grow out of research on atomic electron shells, as actually was the case. In this order of presentation, the energy state diagrams of nuclei and their bearing on γ-radiation are evident from analogy with the phenomena in atomic shells which were treated before. The same is true for the bremsstrahlung of fast electrons and for a number of other processes. In Chap. 6 molecular physics is discussed. It seems unduly far separated from the corresponding treatment of the atoms in Chap. 3; but again historical as well as didactic reasons favor the discussion of molecular structure after that of quantum mechanics to which frequent reference has to be made. Chapter 7, finally, gives a survey of the multitude of phenomena and theoretical explanations in the huge field of solid state physics. From the present state of development, this chapter cannot be as theoretically concise as the other chapters. For this very reason, however, it may give an impression of the variety of problems to be solved, and of the success of the atomistic explanations.

It has already been pointed out in the preface that we do not intend to discuss experimental details. The same applies to the continuously growing number of applications, which are mentioned but not discussed in any detail. Rather it has been the intention of the author to present as vivid a picture of the whole field of atomic physics as possible by stressing the fundamental concepts and the interrelations between the different subjects. Our presentation of atomic physics hopes to lead the reader up to the very limits of our knowledge. By pointing to moot questions and discussing the present trends of active research in our field, the reader should be stimulated to begin thinking independently. He should also realize that atomic physics does not consist of definitely known facts and securely established theories, but that it is rather a very much alive field of fascinating scientific endeavor leading to a better understanding of nature.

Literature

Applications of Atomic Physics

Seeliger, R.: Angewandte Atomphysik, Springer-Verlag, Berlin, 1938.
All textbooks on gas discharges and electronics.
All literature on atomic energy.

Applications to Astrophysics

Chandrasekhar, S.: Principles of Stellar Dynamics, University of Chicago Press, Chicago, 1942.
Jordan, P.: Die Herkunft der Sterne, Wissenschaftliche Verlagsgesellschaft, Stuttgart, 1947.
Rosseland, S.: Theoretical Astrophysics, Oxford University Press, New York, 1936.
Unsöld, A.: Physik der Sternatmosphaeren, Springer-Verlag, Berlin, 1938.

Applications to Biology

Jordan, P.: Physik und das Geheimnis des organischen Lebens, Vieweg, Brunswick, 1941.
Jordan, P.: Eiweissmolekuele, Wissenschaftliche Verlagsgesellschaft, Stuttgart, 1948.

Philosphic Consequences of Atomic Physics

Bavink, B.: Ergebnisse und Probleme der Naturwissenschaften, 8th ed., S. Hirzel, Leipzig, 1945.
Bridgman, P. W.: The Nature of Physical Theory, Princeton University Press, Princeton, N.J., 1936.
Eddington, A. S.: The Philosophy of Physical Science, The Macmillan Company, New York, 1939.

Frank, P.: Between Physics and Philosophy, Harvard University Press, Cambridge, Mass., 1941.
Heisenberg, W.: Wandlungen in den Grundlagen der Naturwissenschaften, 8th ed., S. Hirzel, Leipzig, 1949.
Hermann, G.: Die Bedeutung der modernen Physik fuer die Theorie der Erkenntnis, S. Hirzel, Leipzig, 1937.
Jeans, J. H.: Physics and Philosophy, Cambridge University Press, London, 1943.
Jordan, P.: Physics of the 20th Century, Philosophical Library, New York, 1944.
March, A.: Natur und Naturerkenntnis, Springer-Verlag, Vienna, 1948.
Margenau, H.: The Nature of Physical Reality, McGraw-Hill Book Company, Inc., New York, 1950.
Planck, M.: Wege zur physikalischen Erkenntnis, 3d ed., S. Hirzel, Leipzig, 1943.
Schroedinger, E.: Science and the Human Temperament, W. W. Norton & Company, New York, 1935.
von Weizsäcker, C. F.: Zum Weltbild der Physik, 3d ed., S. Hirzel, Leipzig, 1945.

CHAPTER 2

ATOMS, IONS, ELECTRONS, ATOMIC NUCLEI, PHOTONS

2-1. Proof of the Atomic Structure of Matter and Electricity

It is appropriate to begin a discussion of atomic physics by demonstrating the evidence for the existence of individual atoms and their constituent parts: electrons, ions, and atomic nuclei. For it is the properties and probable behavior of these particles with which atomic physics deals, and from which atomic physicists hope to derive all other properties of matter. When we speak of the atom as being the smallest particle of matter—when we say, for instance, that the iron atom is the smallest particle into which a piece of iron can be divided—we mean that it is the smallest particle of iron that still retains the characteristics typical of the element. We make this distinction because by now it has been established that the atom itself can be broken down into several particles.

It has been only in the present century that experiments have been performed which provide conclusive evidence for the existence of the atom. As recently as in the latter part of the last century a vigorous scientific debate was waged over the question whether the atom existed as an actual physical entity or whether the atomic hypothesis was merely a convenient working hypothesis for explaining many observations which had been made on the behavior of matter. In the following pages we shall discuss the proof for the atomic structure of matter and, at the same time, the proof for the atomicity of the electric charge, and thus of electricity itself, since it is now known that matter and electricity are inseparably related.

The question of the structure of matter, whether it is homogeneous or atomistic, was considered many centuries ago. However, it was not until about 1800, when Dalton conducted his investigations of the structure of chemical compounds, that the atomistic structure of our material world was clearly indicated. Dalton found that in a chemical compound the relative weights of the constituent materials, the elements, are always constant (law of definite composition); and that when two elements can unite with varying weights, these equivalent weights are always integral multiples of the smallest equivalent weight (law of multiple proportions), for example, the weights of oxygen in N_2O, NO,

N_2O_3, NO_2, and N_2O_5 are as 1:2:3:4:5. Dalton's laws are very surprising and difficult to understand if matter is assumed to be homogeneous and arbitrarily divisible. However, from the viewpoint of atomic physics these laws are self-evident, for they are the simple consequence of the fact that always the same number of atoms of the various constituent elements unite to form a particular compound. That is, the ratio of oxygen atoms to nitrogen atoms is always one, two, three, four, or five to two. There are no fractions of atoms.

The concept of the kinetic theory of heat by Krönig and Clausius and its further development, especially by Maxwell and Boltzmann, in the second half of the nineteenth century, brought to light completely independent but equally definite evidence of the existence of individual atoms and of molecules built of atoms. The explanation that gas pressure and its increase with temperature is the result of the collisions of gas atoms or molecules, and that heat conduction as well as internal friction of a gas are the result of transfer of energy and momentum by collisions of the gas atoms or molecules, was such a convincing proof of the kinetic theory of gases and thereby of the existence of individual atoms and molecules that the theory of atomic structure of matter received new support. This evidence for the kinetic theory of gases was especially convincing because, just a decade before, the discovery of Brownian molecular motion had produced very clear experimental confirmation of the temperature dependence of random thermal motion of atoms and molecules.

However, only the present century provided the *direct* and absolutely conclusive proof for the atomic structure of matter, and this proof follows from all the investigations to be discussed in this book. Especially the experiments with canal rays definitely proved that these positive rays consisted of single, charged atoms. The investigation of traces of single atoms or ions in the Wilson cloud chamber (page 251), the discovery of the diffraction of X-rays by the regularly spaced atoms in a crystal which served as diffraction gratings, and, very definitely, the large number of spectroscopic experiments removed the slightest doubt of the existence of the atom. From all these proofs we are convinced that *the atom is the smallest particle into which matter can be divided by chemical methods; whereas by the methods of physics the atom can be divided into its constituent parts which, however, no longer have the characteristics of the atom itself.*

If we assume the existence of the atom, we can then regard Faraday's law of electrolysis, discovered in 1833, as proof of the existence of an elementary quantum of electricity, or, so to speak, of an atom of electricity. That is, if the quantity of a substance transported in an electrolyte depends only on the quantity of charge transported, then each atom migrating as an ion carries the same charge. And further, if, in the case of a monovalent compound, the quantity of the substance deposited by a unit quantity of electricity is proportional to the atomic weight, we must conclude that atoms of equal mass carry equal charges. If a bivalent electrolyte transports the same quantity of electricity and only half as much of the substance is deposited, we must conclude that the bivalent atom always carries two units of electric charge. The existence of an elementary electric charge, designated as e, though appearing extremely probable from Faraday's laws, was not yet definitely proved. This direct proof was provided by the discovery and investigation of the elementary electric charge, called the electron. We shall discuss it in detail on page 28.

2-2. Mass, Size, and Number of Atoms. The Periodic System of Elements

a. Atomic Weights and the Periodic Table

Now that we have discussed the empirical proof of the existence of individual atoms, we shall deal with their properties, especially their masses and diameters. In order to determine the mass of the atom in absolute units, we must know the number of atoms in a mole, Avogadro's number.

The term atomic mass as used in physics or chemistry usually refers not to the absolute value, but to the mass of an atom relative to the mass of an oxygen atom which is arbitrarily taken to be 16.000. This ratio is called the atomic weight. This misnomer is due to a confusion in the conception of mass and weight when science was in its infancy. In this scale hydrogen has an atomic weight of 1.008 and the uranium atom, which until the recent discovery of "transuranium" elements was the heaviest atom, has an average atomic weight of 238.07. The reason for speaking of the average atomic weight will soon be made clear.

Avogadro's law is used to determine the atomic weights of gases. This law states that *under the same conditions of temperature and pressure equal volumes of all gases contain the same number of atoms or molecules.*

Since chemists have shown that the inert gases are monatomic, but that hydrogen, nitrogen, and oxygen are diatomic, the atomic weights of various gases can be determined by weighing them under the same conditions and comparing their weights with that of oxygen. In order to determine the atomic weight of the nongases, it is necessary to form compounds with gases of known atomic weight. If the chemical formula of the compound is known, e.g., Fe_2O_3, the relative atomic weight of the solid element can be computed by measuring the exact mass of the compound and of the residue after completely decomposing the compound and liberating all the oxygen. Relative atomic weights have been measured in this way with the highest precision, especially by Hönigschmid, and are published annually in the international table of atomic weights (see Table 3). The determination of the relative atomic weights provided the basis for the periodic system of elements of Lothar Mayer and Mendelejeff (1869). In this system (see Table 1) the elements are arranged in a sequence of increasing order of atomic weights to form periods, and the periods are arranged to form columns, or groups, so that elements having similar chemical properties such as the alkali metals, the halogens, or the inert gases, each form a group. Thus the monovalent elements are in the first group, the bivalent elements in the second group, and the chemically inactive, inert gases in the eighth and last group of the system. On the left side of the periodic table we have the electropositive elements, so called because they are charged positively in an electrolyte, while in the next to the last group on the right we find the halogens which are very electronegative. The periodic table breaks off with no apparent reason in the middle of a period with uranium, the heaviest element found in nature. It has been one of the important tasks of atomic physics to explain the chemical and physical properties of the elements of the periodic table, only briefly indicated here, as well as its whole arrangement on the basis of the internal structure of the atoms.

It has already been mentioned that the sequence of elements in the periodic table corresponds in general to increasing atomic weight. In several cases, however, the chemical properties of an element make it necessary to deviate from this rule. For example, tellurium with an atomic weight of 127.61 has properties which definitely place it under selenium and in *front* of iodine of atomic weight 126.92, which in turn quite clearly belongs with the other halogens and thus below bromine. Therefore, some departures from the sequence of atomic weights are necessary. Furthermore, it was evident from the periodic table that some elements had not yet been discovered. Between radium emanation (radon) of atomic weight 222.04, which definitely belongs to the inert gases, and radium of the next highest atomic weight 226.05, which just as definitely belongs to the alkaline earths, an element belonging in the alkali metal row was missing. The periodic table thus gives information not only about the known but also about the unknown elements, showing, in addition to their number, their types and properties from the location of the gaps in the table. Many elements were later identified by looking for these predetermined properties. An example is rhenium, which was discovered by Walter and Ida Noddack. Since the atomic weight evidently does not determine the position of an element in the periodic table, the elements have been arranged according to their chemical behavior, and then simply numbered successively beginning with hydrogen and ending with uranium. The numbers assigned to the elements of the periodic table are called their "atomic numbers." We shall soon see that these atomic numbers have a very important physical significance.

An inspection of the relative atomic masses of the elements reveals that in a surprising number of cases they are very close to integral numbers when referred to oxygen equal to 16. However, there are a number of outstanding exceptions to this rule. Therefore Prout, shortly after 1800, proposed the hypothesis (which for a long time lay in oblivion) that all elements might be built up of hydrogen atoms. This idea of the construction of all elements from *one* basic substance received new support when at the beginning of this century the study of radioactive elements (page 273) led to the discovery of isotopes. It was learned that there are atoms which have different masses in spite of the fact that they belong to the same element and thus have the same atomic number and chemical properties. Such atoms are called *isotopes* and the atomic masses of these isotopes are actually close to being integral numbers. The outstanding nonintegral atomic weights, such as chlorine of atomic weight of 35.457, arise from the mixture of several isotopes of different masses. In the case of chlorine it is the mixture of the isotopes of weights 35 and 37. We shall go into the important phenomenon of isotopes more fully on page 43.

If we consider molecules instead of atoms, we must likewise think of the molecular weight instead of the atomic weight. If we know the chemical formula, we can compute the molecular weight from the weight of the constituent atoms; for example, the molecular weight of methane, CH_4, is, in round numbers, $12 + 4 \times 1 = 16$. The number of grams of a substance equal to the atomic or molecular weight of the substance is called a mole. One mole of molecular hydrogen, H_2, is 2 gm; 1 mole of water, H_2O, 18 gm of water; 1 mole Hg, 200.6 gm of mercury. By applying the mole concept according to Avogadro's law, which states that equal volumes of gases under the same external conditions have equal numbers of atoms or molecules, to substances other than gases, we see that a mole of any substance has the same number of atoms or molecules as a mole of any other substance. Thus, we can compute the absolute mass of an atom or molecule from the relative atomic mass if we know the number of atoms in a mole, the so-called

Avogadro's number N_0. This was determined in 1865 by Loschmidt to be about 10^{23}.

b. The Determination of Avogadro's Number and the Absolute Mass of the Atom

At present there are many completely different methods by which Avogadro's number can be measured. The fact that there is excellent agreement in the results is the best proof of its correctness.

The method used by Loschmidt made use of the expressions obtained from the kinetic theory of gases for the viscosity of a gas,

$$\eta = \frac{M\bar{v}}{16\pi r^2 N_0} \tag{2-1}$$

and for the coefficient of heat conductivity

$$k = \frac{c_v\bar{v}}{16\pi r^2 N_0} \tag{2-2}$$

Here M is the molecular weight, r the radius of the molecule as determined by the kinetic theory—we shall consider this in greater detail when we discuss the size of the atom—c_v, the specific heat at constant volume, $\bar{v}$ the average velocity of the molecules due to their thermal motion, and N_0 Avogadro's number. Since η and k can be measured experimentally and r follows from the constant b in the van der Waals equation of state (2-15), these two equations give us a means of computing N_0. Because of the difficulties encountered in measuring the values of r, η, and k accurately, this method is not very exact.

The same difficulties are encountered in many other possible ways of determining this important fundamental constant of atomic physics. Two of the methods depend on measuring the Boltzmann constant k which is equal to the quotient of the general gas constant R and Avogadro's number,

$$k = \frac{R}{N_0} \tag{2-3}$$

To measure k, Perrin (1909) made use of the fact that minute suspended particles tend to settle out of a liquid, but that by diffusion they tend to maintain a uniform density. The decrease in density with the height of the liquid corresponds to the decrease of the density of an isothermal atmosphere with height. According to the kinetic theory of gases the decrease of the particle density between the height 0 and h is given by the formula

$$n_h = n_0 e^{-mgh/kT} \tag{2-4}$$

Table 1. The Periodic Table of Elements

	I	II	III	IV	V	VI	VII	VIII			
1	**1** H 1.008										**2** He 4.003
2	**3** Li 6.940	**4** Be 9.02	**5** B 10.82	**6** C 12.01	**7** N 14.008	**8** O 16.0000	**9** F 19.00				**10** Ne 20.183
3	**11** Na 22.997	**12** Mg 24.32	**13** Al 26.97	**14** Si 28.06	**15** P 30.98	**16** S 32.06	**17** Cl 35.457				**18** Ar 39.944
4	**19** K 39.096	**20** Ca 40.08	**21** Sc 45.10	**22** Ti 47.90	**23** V 50.95	**24** Cr 52.01	**25** Mn 54.93	**26** Fe 55.85	**27** Co 58.94	**28** Ni 58.69	
4	**29** Cu 63.57	**30** Zn 65.38	**31** Ga 69.72	**32** Ge 72.60	**33** As 74.91	**34** Se 78.96	**35** Br 79.916				**36** Kr 83.7
5	**37** Rb 85.48	**38** Sr 87.63	**39** Y 88.92	**40** Zr 91.22	**41** Nb 92.91	**42** Mo 95.95	**43** Tc 99	**44** Ru 101.7	**45** Rh 102.91	**46** Pd 106.7	
5	**47** Ag 107.88	**48** Cd 112.41	**49** In 114.76	**50** Sn 118.70	**51** Sb 121.76	**52** Te 127.61	**53** I 126.92				**54** Xe 131.3
6	**55** Cs 132.91	**56** Ba 137.36	**57** La 138.92 **58–71**	**72** Hf 178.6	**73** Ta 180.88	**74** W 183.92	**75** Re 186.31	**76** Os 190.2	**77** Ir 193.1	**78** Pt 195.23	
6	**79** Au 197.2	**80** Hg 200.61	**81** Tl 204.39	**82** Pb 207.21	**83** Bi 209.00	**84** Po 210	**85** At 211				**86** Rn 222.04
7	**87** Fr 223	**88** Ra 226.05	**89** Ac 227.05 **90–96**								
6	**58** Ce 140.13 **59** Pr 140.92	**60** Nd 144.27 **61** Pm 147	**62** Sm 150.43 **63** Eu 152.0	**64** Gd 156.9 **65** Tb 159.2	**66** Dy 162.46 **67** Ho 164.94	**68** Er 167.2 **69** Tm 169.4	**70** Yb 173.04 **71** Lu 174.99				
7	**90** Th 232.12 **91** Pa 231	**92** U 238.07 **93** Np 237	**94** Pu 239 **95** Am 241	**96** Cm 242 **97** Dk 243	**98** Cf 244						

where n_h, n_0, the mass of the particle, m, and the absolute temperature T can be measured experimentally. Substituting these values in (2-4) the value of k can be calculated.

Closely related to this method of Perrin's is the determination of k [and N_0 from Eq. (2-3)] by measuring the average lateral displacement of suspended particles due to their Brownian motion. The theory, developed first by Einstein in 1905, shows that the mean square displacement $\overline{x^2}$ of spherical particles of radius r_0 moving in a gas of viscosity η at an absolute temperature T after an observation time τ is given by

$$\overline{x^2} = \frac{kT\tau}{3\pi\eta r_0} \tag{2-5}$$

$\overline{x^2}$ can be measured and, since all other constants are known, k can be computed.

The Boltzmann constant k and, by using (2-3), N_0 also can be determined from radiation measurements. According to the Stefan-Boltzmann law, the total radiation from a blackbody at absolute temperature T is given by

$$S = \frac{2\pi^4 k^4}{15c^2h^2} T^4 \tag{2-6}$$

while the energy radiated in the wavelength interval λ to $\lambda + d\lambda$ is given by Planck's law

$$J_\lambda = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/k\lambda T} - 1} \tag{2-7}$$

After measuring S and J the two unknown constants h and k can be computed by using Eqs. (2-6) and 2-7), and then N_0 again follows from Eq. (2-3).

Another optical method makes use of the scattering of light by small particles (Rayleigh scattering) by which the attenuation of the incident radiation depends upon the number of particles (molecules of air) in the scattering medium. The measurement of the attenuation due to Rayleigh scattering thus provides another method for determining N_0.

Of the two most exact methods, one is based on determining the molecular weight M and density δ of the most ideal calcite crystal, $CaCO_3$, it is possible to find. Since the geometrical structure of the crystal can be very exactly determined by the X-ray diffraction method (page 426), it is possible to compute the exact volume of a single molecule in the crystal and from that information obtain Avogadro's number

$$N_0 = \frac{M}{\delta V} \tag{2-8}$$

This method became a precision method after the absolute wavelengths of the X-rays used to determine the lattice constants of the $CaCO_3$ crystal were measured accurately by means of optical gratings.

The second precision method follows from Faraday's laws of electrolysis according to which a quantity of electricity

$$F = 96{,}496 \pm 7 \quad \text{amp-sec} \tag{2-9}$$

is required to deposit one mole of a substance. Then, if we know the magnitude of the elementary electric charge e, which we shall consider later, we can use the relation

$$N_0 = \frac{F}{e} \tag{2-10}$$

to determine N_0. The last two methods give the best value for N_0 that we have at the present time. It is

$$N_0 = (6.0235 \pm 0.0004) \times 10^{23} \tag{2-11}$$

From the relative atomic weight of an atom A based on the value of oxygen = 16 we can obtain the absolute mass of the atom from

$$M_A = \frac{A}{N_0} \tag{2-12}$$

From (2-12) we can calculate the absolute mass of the hydrogen atom to be

$$M_H = 1.672 \times 10^{-24} \quad \text{gm} \tag{2-13}$$

c. *The Size of the Atom*

We now return to the question of the size of the atom wherein we encounter a fundamental difficulty, namely, how to define the radius of the atom. The same problem is met in considering the radius of the electron and of the nucleus. If the radii of atoms could be defined exactly, it would mean, as in the case of two equal billiard balls, that when the balls pass each other at a distance between their centers greater than twice the radius of either, they would pass undeviated. But if the displacement were less than twice the radius, they would collide and each ball would be deviated. The impact radius r and the impact cross section πr^2 can be exactly defined in the case of two billiard balls: the force exerted by one ball on the other falls off very suddenly at a definite radius, i.e., with a high power of $1/r$. If we consider the other extreme case, that of the electrically charged particles, such as electrons and ions, which we shall consider in Sec. 2-4, then as the particles pass each other at any arbitrary distance, theoretically, at least, Coulomb forces of attraction or repulsion are exerted so that it is impossible to speak of a definite impact radius.

The atom, as we shall see in the next section, lies between the two extreme cases: it cannot be regarded as a rigid ball with an exactly defined radius; however, the radius is fairly well defined as the forces which are exerted by the atom in a collision decrease rapidly with increasing distance. If the atomic radius or atomic volume $4/3\pi r^3$ is determined by several methods, somewhat different values will be obtained because the definition of the radius is inherently linked with the method of measurement. It is possible, for instance, to determine the atomic volume, and from it the radius of the atom under consideration, from the density of a crystal or liquid (e.g., a liquefied rare gas). However, we have to keep in mind that the atoms here, in addition to their own volume, require additional space for vibration, even at the absolute zero of temperature (zero-point vibration, page 220). Because of the inexact definition of the radius, the volume required by a single atom and thus its radius also depends upon its environment, that is, on the forces exerted by neighboring atoms.

The radius of an atom can also be determined from the average distance, λ, which it travels in a gas of density N (atoms per cm^3) without colliding with another atom (mean free path), since the coefficient of internal friction of the gas which can be measured is a function of λ. The formula relating the mean free path λ to the radius r of the atom is, according to the kinetic theory of gases,

$$\lambda = \frac{1}{4\sqrt{2}N\pi r^2} \qquad (2\text{-}14)$$

From this formula it is evident that, for a given density N, λ is inversely proportional to the effective cross section πr^2 of the atom. This differs from the case of the billiard balls with fixed radii because van der Waals' forces of attraction act upon the atoms so that the effective cross section depends upon the impact velocity, i.e., the temperature (Sutherland's theory), resulting in what might be called a "softness" of the surface of the atom. The atomic volume can also be determined from the van der Waals equation of state

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT \qquad (2\text{-}15)$$

since the constant b, the co-volume, according to theory is four times the atomic volume.

The atomic radii as determined by all these methods are of the order of magnitude 10^{-8} cm, yet the differences in the results of the various methods exceed considerably the experimental error. As an example of these results and the dependency of the size of the atom on its position in the periodic table, we show in Fig. 1 a plot of the atomic volume of all elements versus atomic number, and in Fig. 2 a plot of atomic radii and the radii of the corresponding ions of the atoms. In both illustrations the periodic structure of the periodic table stands out clearly. The curve of the atomic volume, Fig. 1, was computed by

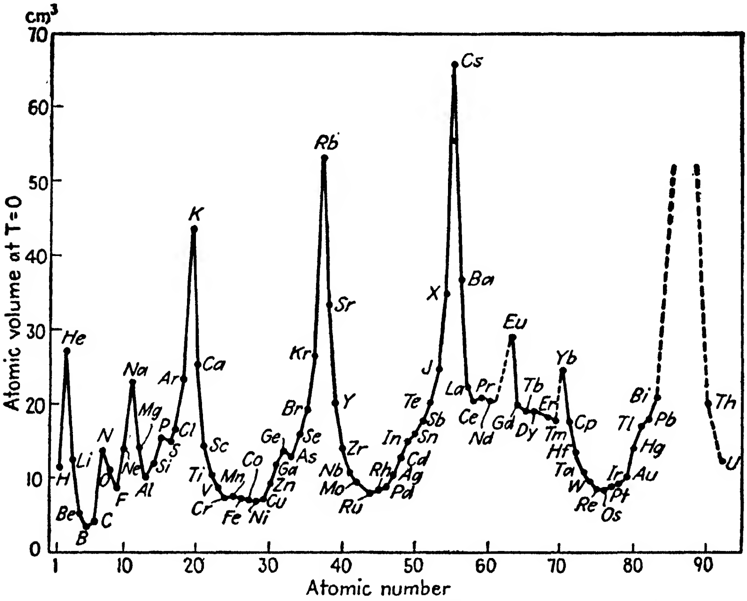


FIG. 1. Atomic volume of the elements plotted against the atomic number for use in calculating atomic radii.

dividing the atomic weight by the density of the element in the condensed state, the density being referred to absolute zero. On the other hand the atomic and ionic radii in Fig. 2 were determined from the distances between atomic centers in different crystal lattices.

2-3. Nuclei and Electron Shells as Structural Elements of the Atoms. Atomic Models

Originally the atom, as its name implies, was assumed to be the smallest particle into which matter could be divided, and it was the general concept that the entire material world was built from about 90 different kinds of atoms. We know now that the atom can be divided, not by chemical methods, but by the methods of physics. Atomic

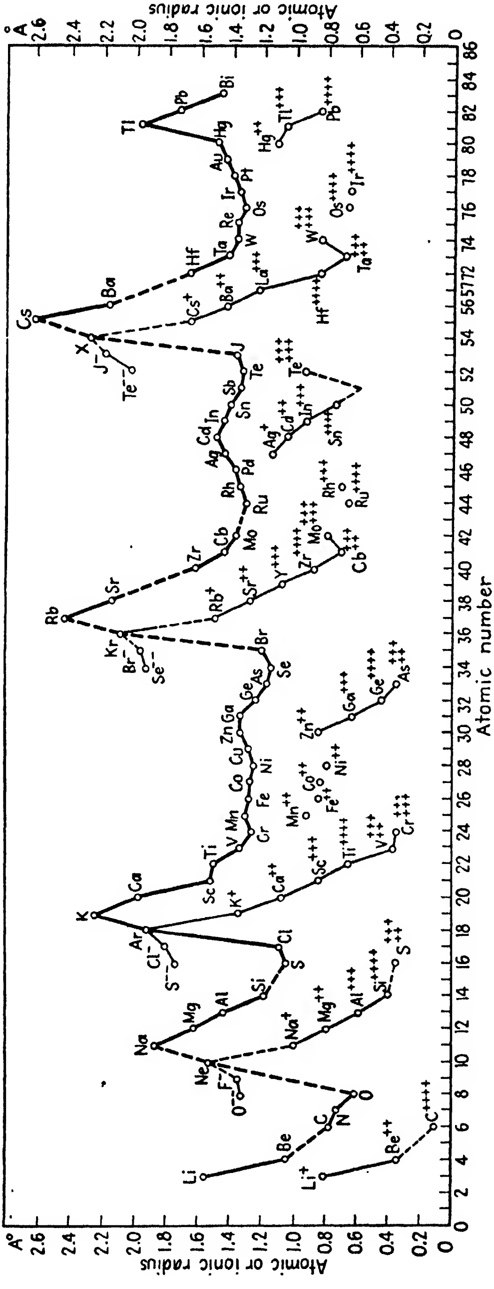


FIG. 2. Radii of atoms and their positive ions determined from interatomic distances in crystal lattices.

physics starts with the realization that the atom is divisible. The first evidence of the structure of the atom came from the physics of gas discharges from which it was evident that neutral atoms and molecules were split up into electrically charged particles, electrons and ions.

The study of atomic structure really began with Lenard's experiments on the passage of cathode rays, or fast electrons, through thin metal foils. Lenard discovered first that fast electrons can penetrate a large number of atoms without noticeable deviation from their original direction. From this he drew the correct conclusion that atoms should not be thought of as massive balls, but that they consist mostly of empty space. Lenard also found that for a given electron velocity the mass of the scattering centers per cubic centimeter determines the amount of scattering. In other words, a few atoms of large mass scatter as much as many atoms of small mass, and for equal masses the scattering power is approximately the same, independent of the state of aggregation. The most important discovery of Lenard was that the atomic radius as determined by scattering depended on the velocity of the impinging electrons. For slow electrons the atomic radius is equivalent to that determined by the kinetic theory, 10^{-8} cm, but with electrons of increasing velocity the effective atomic radius ultimately decreases to about 10^{-12} cm, one ten-thousandth of the original value! As a result of these experiments Lenard came to the conclusion, which has since been confirmed, that the atom consisted of a small massive nucleus surrounded by a force field of 10^{-8} cm radius, and that this force field deflected (scattered) slow electrons but did not affect fast electrons very much. Lenard thought this force field to consist of an

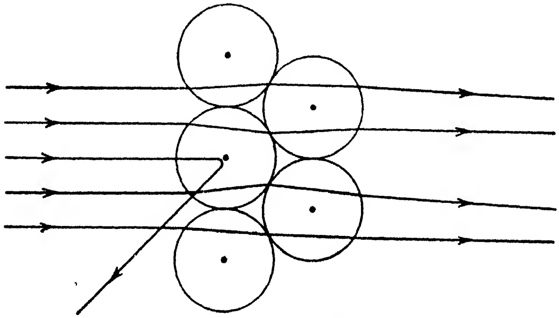


FIG. 3. Schematic representation of the scattering of α-particles by the atoms of a thin metal foil, illustrating small scattering by the electrons of the shells, large deflections only by collision with atomic nuclei.

equal number of positive and negative charges regularly arranged within the atom, making the whole atom electrically neutral.

This new concept of the atom, proposed in 1903, proved to be fairly correct. It was confirmed and extended quantitatively by Rutherford about 1912 by his famous scattering experiments. Instead of electrons Rutherford used α-particles which are about 7,000 times as heavy as electrons and carry a double positive charge. He found that the α-particles could penetrate thousands of atoms without any noticeable deviation and that only very rarely one α-particle would be deflected (scattered) through a very large angle (Fig. 3). A few years later Wilson invented his cloud chamber (page 251), with which he could make directly visible the scattering of single α-particles (Fig. 4). Rutherford concluded that since scattering through a large angle occurred only rarely, the radius of the scattering center of the atom actually was of the order of only 10^{-12} to 10^{-13} cm. Furthermore, almost the whole mass of the atom must be concentrated in the small "nucleus" of the atom, for otherwise the heavy α-particle could not be scattered through such a wide angle. Rutherford thus conceived of the atom as consisting of a positively charged nucleus containing almost all of the mass of the atom, surrounded by almost massless electric charges (electrons). These would rotate around the nucleus with such velocity that the centrifugal forces would compensate the electrostatic attractive forces. According to this conception the α-particles can be scattered only by the nucleus because the very light electrons cannot deflect the heavy α-particles.

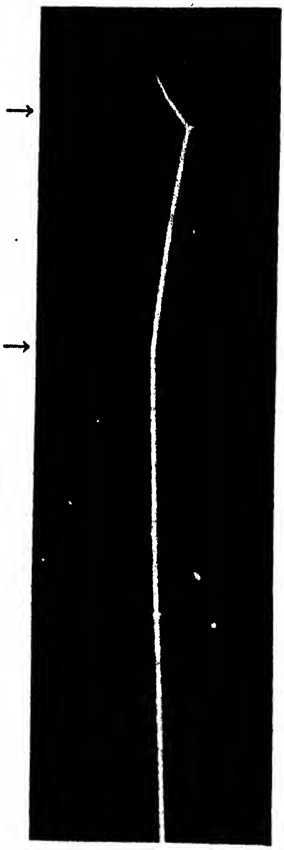


FIG. 4. Cloud-chamber photograph (C. T. R. Wilson): double deflection of an α-particle by collisions with nuclei of air molecules. The short track toward the right from the point of the second deflection is due to the recoil of the deflecting nucleus. In the straight portion between the two deflections (marked by arrows) the α-particle permeates approximately 10^5 gas molecules without perceivable deflection.

To test the correctness of this atomic model Rutherford treated the scattering process theoretically and then compared the theoretical results with his measurements. Since the α-particle of charge $+2e$ is scattered by the nucleus of charge $+Ze$, there acts between the two particles a Coulomb force (Fig. 5)

$$F = \frac{2Ze^2}{r^2} \tag{2-16}$$

which depends on r in the same manner as does the gravitational force, derived from planetary motion, mM/r^2. Furthermore, the orbit of the α-particle must be a conic section with the nucleus at a focal point, and because of the repulsive force between the two like-charged bodies, the orbit is a hyperbola as in Fig. 5, degenerating into a straight line for a central collision. Let p be the perpendicular distance between the nucleus and the original direction of the α-particle (impact parameter),

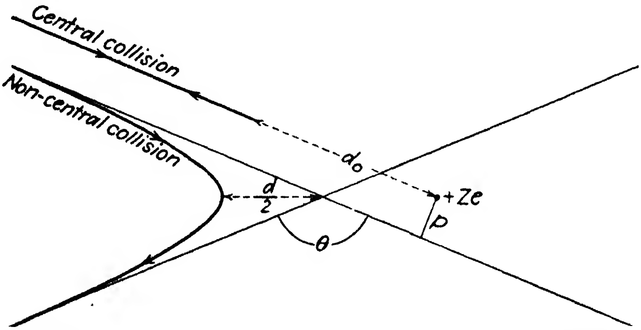


FIG. 5. Central and noncentral collision between α-particle and nucleus of charge $+Ze$.

d_0 be the closest distance to which the α-particle with a velocity v in a central collision actually approaches the nucleus, and ϑ the angle of deflection, which becomes 180° for a central collision. Then the value of d_0 follows from the condition that at this distance all the initial kinetic energy has been transformed into potential energy, so that we have

$$\frac{m}{2}v^2 = \frac{2Ze^2}{d_0} \qquad d_0 = \frac{4Ze^2}{mv^2} \tag{2-17}$$

It follows from the laws of conservation of energy and angular momentum that this minimum distance d_0 is equal to the distance between the vertices of the two branches of the hyperbola. Since, for the degenerated hyperbola, the focal point and vertex coincide, this distance d_0 is reached only in a central collision. From Fig. 5 we get the deflection angle ϑ as a function of the impact parameter p,

$$\cot\frac{\vartheta}{2} = \frac{2p}{d_0} = \frac{pmv^2}{2Ze^2} \tag{2-18}$$

Since the number of α-particles scattered into a definite angle can be measured, (2-18) can be used to calculate the fraction dn/n of α-particles scattered through the angle ϑ into the solid angle $d\Omega$ by a foil of thickness d with N atoms per cm³ each with a positive charge $+Ze$. By assuming that in a thin foil an α-particle is scattered only once, Rutherford's formula

$$\frac{dn}{n} = \frac{2dNZ^2e^4}{m^2v^4} \frac{1}{\sin^4(\vartheta/2)} d\Omega \qquad (2\text{-}19)$$

can be computed by statistical methods. The theoretical dependency of the scattering on $1/v^4$ and $1/\sin^4(\vartheta/2)$ was shown to be in agreement with experiment by exact measurements, especially by Geiger. The slight deviation from this formula, especially for the scattering of slow and heavy particles, was explained meanwhile by Gerthsen, who applied wave mechanics, which we shall consider later (page 206), and arrived at the correct scattering formula. Rutherford's formula and his atomic model from which the formula was derived thus have been proved essentially correct.

The question of the absolute magnitude of the positive charge Z of the scattering nuclei was left open, for it was at first assumed that it was approximately equal to half of the atomic weight A of the scattering atom. This question was answered by Chadwick in 1920 when, by simultaneously measuring dn and n, he determined the absolute number of charges Z of the scattering nuclei and thereby established the important law that the charge on the nucleus Z is exactly equal to its atomic number N in the periodic table. Because the atom is electrically neutral, the number of electrons encircling the nucleus must be equal to the positive charges on the nucleus. Chadwick's law thus may be stated

Atomic number N = number of positive charges Z of the
nucleus = number of electrons of the atom

The atomic number which at first was a purely formal basis for numbering the elements in the periodic table had gained a very decisive physical significance for the structure of the atom. Meanwhile Chadwick's results with α-rays have been confirmed by measurements with fast cathode rays.

Moseley, about 1913, in an entirely different way obtained essentially the same results. He investigated the X-ray spectra of all available elements, and found that the wavelength of the shortwave principal line, K_α, became shorter in a regular manner in going from the light to the heavy elements (Fig. 6). The frequency $\nu = c/\lambda$ of the K_α lines of all elements (atoms) can be determined quantitatively from the formula (Fig. 7)

$$\nu = \text{const } (N-1)^2 \tag{2-20}$$

where N increases by one in going from one element to the next in the periodic table, and was identified with the atomic number N of the element. Thus Moseley had for the first time found a purely physical method for determining the atomic numbers of the elements and could justify the arrangement of tellurium and iodine in the periodic table which we discussed on page 13 as contradictory to the steady increase

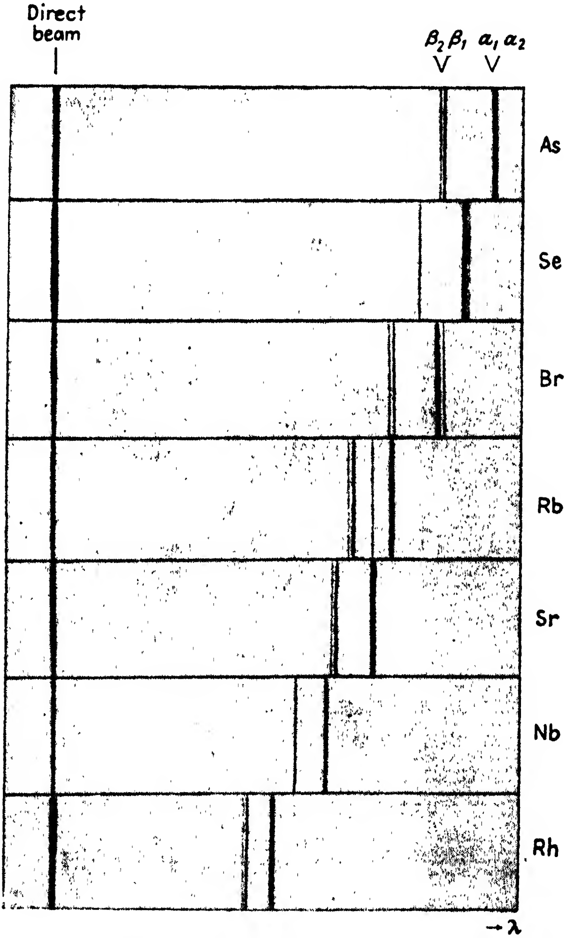


FIG. 6. Shortwave X-ray spectra (K lines) of the elements As^{33}, Se^{34}, Br^{35}, Rb^{37}, Sr^{38}, Cb^{41}, and Rh^{45} illustrate the increase of the frequency with increasing atomic number (Moseley's law). (*After Siegbahn.*)

of the atomic weight. A little later van den Broeck showed that Moseley's atomic number N was identical with the charge on the nucleus Z, a conclusion later confirmed by Chadwick's scattering experiment about which we have already spoken.

From the experiments described in this section it became clear that an atom, in contrast to the old, naïve concept of a homogeneous particle of impenetrable matter, actually consists mostly of empty space with a very small, but extremely heavy, positively charged nucleus surrounded by a shell of almost massless negative electrons, held together as a system by electrical forces. Also it was made clear that the periodic table is more than a practical table of, by now, 96 elements for chemists, but that it is very closely related to the structure of the atom. In particular, the atomic number of the atom in the periodic table, because it is the number of positive charges of the nucleus and identical with the number of electrons in the atomic shell, turns out to be one of the most important figures for the structure of the atom.

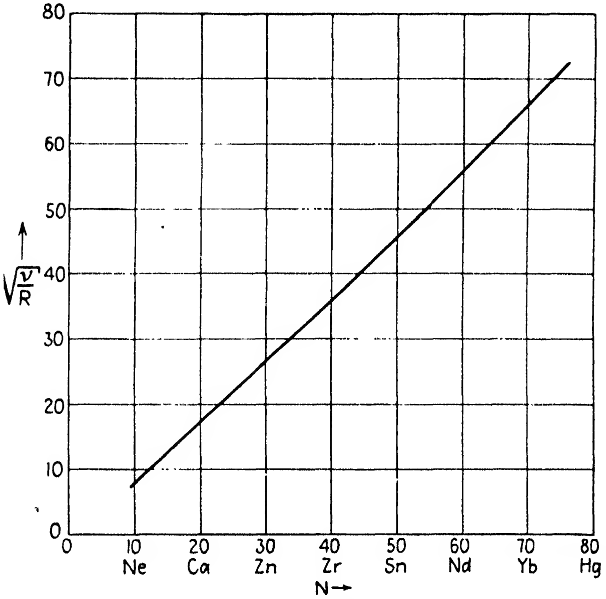


FIG. 7. The approximately linear increase of the frequency of the X-ray line K_{α_1} with the square of the atomic number of the elements (Moseley curve).

The pictorial concepts of the atom, derived from experimental results in the way discussed above and to be continued later, are called *atomic models.* The various models are improved and developed constantly by new research results, and it is our belief that this process of continuous improvement of the atomic model leads to an increasingly better agreement of the properties of the theoretical model with those of the actual atom. Lenard started with the old model of the atom based on chemical knowledge and the kinetic theory and then improved it by using the results of his scattering experiments, thus arriving at a model consisting of positive and negative charges held together by electrical forces with a nucleus one ten-thousandth the diameter of the atom. Rutherford extended the model of Lenard by showing that practically all the mass and positive charge of the atom was concentrated in the nucleus and that the encircling electrons were held in place by equilibrium between centrifugal and electrostatic forces, thus arriving at a dynamically but not statically stable atom. We shall deal with the further extension of this model by Bohr, Sommerfeld, Heisenberg, and Schrödinger in the following chapters.

2-4. Free Electrons and Ions

In the last section we considered the evidence for the discontinuous structure of the atom and learned that it consists of positive and negative charges. We also learned that the electron, the name given to the elementary negative charge in the atom, had very little mass, and that most of the mass of the atom was concentrated in the nucleus. In this section we shall be concerned with the production and characteristics of free electrons and ions (as atoms or molecules with an excess negative or positive charge are designated).

Electrically charged atoms, ions, were first observed in electrolysis. When sodium chloride is in solution and an electric field is applied, the sodium atoms migrate to the negative pole and the chlorine atoms to the positive pole, demonstrating that they must be positively and negatively charged respectively. We speak of Na^+ and Cl^- ions; and we know now the process of their formation. The electropositive sodium atom, whose place is in the first group of the periodic table, gives up an electron to the chlorine atom which is in the seventh group. In this case an electron jumps from one atom to another, but it does not turn up as a free electron. Quite generally the liquid state is not well suited for the study of free ions or electrons because of the mutual interaction of the closely packed atoms or ions, respectively. In the following discussions we therefore deal with gases, in which electrons and ions can be considered reasonably free.

a. The Production of Free Electrons

We begin with a treatment of the electron. But before we concern ourselves with its characteristics and mass we shall examine the possibility of producing free electrons. The most important methods of producing free electrons are: *the release of electrons by impact ionization of gas atoms, especially in electric discharges; the "evaporation of electrons" from an incandescent metal surface; and the release of electrons from atoms or solid surfaces by irradiation with shortwave light, the photoelectric effect.*

Ionization is the process of separation of an electron from a neutral atom or molecule, which then becomes a positive ion, or, under certain conditions, from an ion, which then becomes a doubly charged ion. Ionization by impact is the separation of an electron by collision with a fast particle, whether it be a radioactive particle, an electron or ion accelerated by an electric field, or an electron, ion, or neutral atom with random high velocity due to sufficiently high temperatures. Impact ionization by radioactive particles serves as a method of detection of such particles in a cloud chamber, ionization chamber, or counter tube, as we shall discuss in detail when considering the experimental methods of nuclear research (page 249). Ionization by impact in a gas at high temperatures (together with the simultaneous photoionization to be discussed on page 106) is called *thermal ionization.* The processes of thermal ionization together with the counteracting formation of neutral atoms by recombination of ions and electrons (page 105) yields an exponentially increasing degree of ionization α, dependent on the temperature. α can be calculated from *Saha's equation*

$$\frac{\alpha^2}{1 - \alpha^2} p = 3.2 \times 10^{-7} T^{2.5} \times 10^{-5{,}039\, E_i/T} \qquad (2\text{-}21)$$

where p = pressure and E_i = ionization potential of the atoms in volts. Thermal ionization plays an essential role only at temperatures of several thousand degrees Kelvin as in arc discharges, in the hottest flames, and in the atmosphere as well as in the inaccessible interior of the fixed stars. In the high-current carbon arc, for instance, it is possible to produce at atmospheric pressures above 10^{17} free electrons per cubic centimeter so that on the average every fifth atom has lost an electron. The impact collision of electrons accelerated by an electric field is important for the mechanism of glow discharge and of spark breakdown (also for lightning). Ions are too heavy and large to be important in this respect. These discharges require primary electrons produced by ionization by radioactive rays or by photoelectric release from one of the electrodes. Electrons, accelerated by a very strong field in a vacuum discharge (10^{-6} atmosphere) to very high velocities, are called cathode rays (Goldstein, 1876). According to Lenard, cathode rays can penetrate thin aluminum foils and leave the discharge tube to enter free space (Lenard window) and there be investigated. We have already discussed Lenard's scattering experiments with cathode rays on page 22. Related to ionization by impact of gas atoms is the release of electrons from a metal by positive ions impinging on the cathode of a glow-discharge tube, or by fast electrons, the effect of secondary electron emission. We shall return to the consideration of these phenomena on page 474.

To ionize an atom a certain amount of energy, the ionization energy, must be expended. The energy required varies from atom to atom and depends on the binding force between the rest atom and the electron to be released. Table 2 contains the ionization energies of some of the more important atoms and molecules. The ionization energy for impact ionization is supplied by the kinetic energy of the colliding electrons. The energy can be measured in ergs, but usually it is given in terms of the accelerating potential in volts. The kinetic energy of an electron, which is accelerated by the potential V, is given by the equation

$$eV = \frac{m}{2}v^2 \tag{2-22}$$

In this sense it is customary to speak of the ionization potential in volts or, since charge times potential has the dimension of energy, of the ionization energy in eV = electron volts. For computing the energy in ergs and other energy units see (3-10) and (3-11), page 86.

Table 2. Ionization Energies (in Electron Volts) of Some Important Atoms and Molecules

Gas or vapor	*Ionization potential* (ev)
H	13.6
H_2	15.4
C	11.2
N	14.5
N_2	15.8
O	13.6
O_2	11.2
He	24.5
Ne	21.5
Ar	15.7
Hg	10.4
Na	5.1
Cs	3.9

In the second method for producing free electrons, the photoelectric or photoeffect, the energy is supplied by the incident radiation. Lenard's discovery that the kinetic energy of the released electrons is proportional to the frequency of the incident light was essentially the beginning of the quantum theory (page 64). For the quantum theory states that the energy of light is proportional to its frequency (i.e., inversely proportional to its wavelength), whereas according to the classical theory the energy of any radiation is proportional to its intensity. It is, consequently, in agreement with quantum theory, if the wavelength of light necessary to free electrons is the shorter, the larger the binding energy of the electron to its atom or metal surface. Only the *number* of electrons released per second per square centimeter is dependent upon the *intensity* of the radiation. Photoelectric electrons can be released from isolated atoms and molecules as well as from large atom complexes, especially metal surfaces. In the first case, to which we shall return on page 106, we speak of photoionization of the atom or molecule. Only in the second case do we use the term photoeffect in its restricted sense. In addition to this "external" photoeffect there exists the closely related phenomenon of the release of electrons by absorption of radiation in the interior of a crystal, the "internal" or semiconductor photoeffect, which we shall consider further in connection with solid state physics, page 467. A practical application of the external photoeffect is the photoelectric cell. According to Fig. 8 it consists of a metallic layer on the inside of an evacuated glass vessel. Electrons are released from the vacuum-evaporated metal layer F by the incident radiation and collected by the anode A. The photocurrent can be measured directly by a galvanometer, by an amplifier, or by measuring the voltage drop in a high resistance R with an electrometer. This photoelectric cell has found wide use as receiver of radiation.

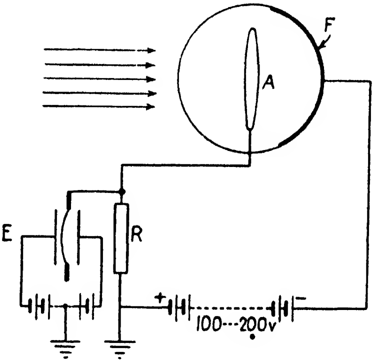


FIG. 8. Photoelectric cell with circuit for electrometric measuring. F = photosensitive alkali layer, A = anode, R = high resistance.

The most important method for producing free electrons utilizes a phenomenon discovered by Elster and Geitel in 1882, the evaporation of electrons from an incandescent metal. We shall discuss the atomistic mechanism of this thermionic emission on page 446. The number of emitted electrons increases exponentially with increasing temperature. It depends furthermore on the so-called work function, i.e., the energy (related to the ionization energy) which is necessary to overcome the binding force and to release an electron from the metal. For a metal with a work function ϕ, the saturation current of electrons leaving a metal surface is given by Richardson's equation

$$i = AT^2 e^{-\phi/kT} \quad \text{amp/cm}^2 \tag{2-23}$$

where A is a constant of the metal. Vacuum tubes, X-ray tubes, and nearly all technical electronic instruments utilize this emission of electrons from electrically heated wires, the so-called thermionic cathodes. A number of finer details of this effect will be treated on page 447.

b. The Measurement of the Charge and Mass of the Electron

Now that we have learned how free electrons are produced we shall take up the properties of this important constituent elementary particle of the atom. We begin with the mass of the electron and its charge, which we have already defined to be equal to the elementary quantum of electricity e.

There are numerous methods for measuring the charge of the electron. From Avogadro's number N_0 and Faraday's constant F (the quantity of electricity required to deposit one mole of a monovalent substance) it follows that the charge on the electron is

$$e = \frac{F}{N_0} = 1.60 \times 10^{-19} \quad \text{amp-sec (coulomb)} = 4.80 \times 10^{-10} \quad \text{esu} \tag{2-24}$$

Here the accuracy of e depends upon the accuracy with which F and N_0 can be measured.

Historically e was first determined in a very similar manner, by measuring the total charge transported by a definite number of charge carriers; for example, by counting a large number of α-particles and measuring with an electrometer the charge transported by them. The neatest and most direct method, however, is the one devised by Millikan to measure the charge of a single electron. In this, the famous oil-drop method, the motion of microscopic oil drops between the plates of a condenser (Fig. 9) was observed after electrons, which had been produced photoelectrically in a gas, attached themselves to the droplets. If the upper plate is charged positively and the field between the plates is $E = V/d$, where V is the potential difference and d the separation, then the forces acting on the droplet are: the downward force of gravity, mg, and the oppositely directed electrostatic force between the charged plates, eE. If the field in the condenser is adjusted so that the droplet is held stationary between the plates, the charge e can be directly computed from the equilibrium condition

$$eE = mg \tag{2-25}$$

The weight of the droplet can be calculated by substituting its radius r, measured micrometrically, and the densities of the oil δ and air

δ_a (to account for the buoyancy of the oil drop in air) in the expression

$$mg = \frac{4\pi r^3(\delta - \delta_a)g}{3} \tag{2-26}$$

Actually, however, the droplet cannot be held in equilibrium accurately enough nor can the radius be measured exactly. Consequently, the vertical fall or acceleration of the droplet with and without a charge on the plates must be measured. With no field the droplet falls with a constant velocity resulting from the equilibrium between the force of gravity and the force of friction of air (viscosity η), which can be computed from Stokes' law which we will assume,

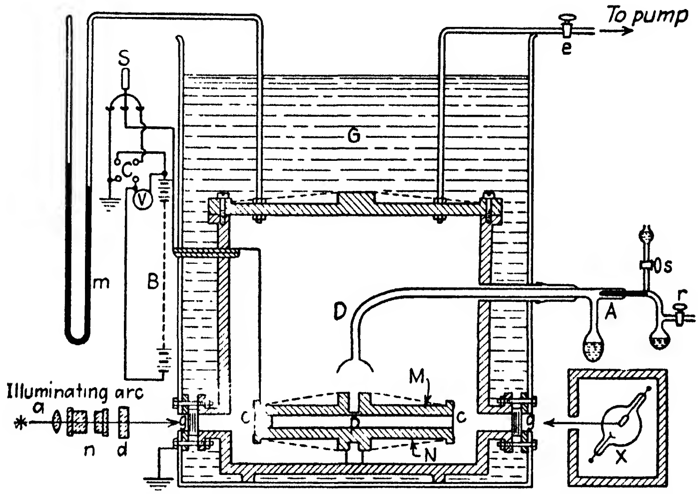


FIG. 9. Representation (partly schematic) of Millikan's oil-drop method for measuring the charge e of the electron. p = plate condenser.

$$v_0 = \frac{mg}{6\pi\eta r} = \frac{2r^2 g(\delta - \delta_a)}{9\eta} \tag{2-27}$$

Since all quantities in (2-27) except r are known, it can be computed by measuring the velocity of fall of the droplet with no charge on the condenser. If now the condenser is charged so that a field E exists between the plates, there is an additional force acting on the droplet, eE, whose direction depends on the polarity of the plates. Under these conditions the droplet falls (or rises) with a different velocity given by

$$6\pi\eta r v_1 = \frac{4\pi}{3} r^3(\delta - \delta_a)g \pm eE \tag{2-28}$$

Since now the radius r is known from measuring the velocity of the same droplet falling in a field-free space (2-27), only the charge e is unknown in (2-28). After making measurements on many droplets Millikan found that the preponderant value of e was that given in (2-24). In some cases the charge appeared twice or three times as large because more than one electron attached itself to one droplet. Thus the existence of an elementary quantum of electricity was demonstrated by direct measurement of the charge of single particles. This value of e agreed within the limits of experimental error with the method used in obtaining the value given in (2-24).

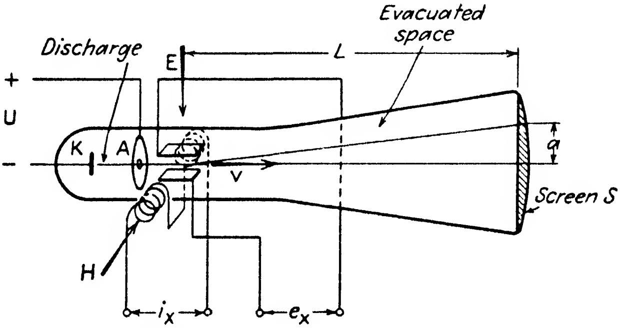


FIG. 10. Cathode-ray oscilloscope (Braun tube) with electric and magnetic deflection of the beam for measuring the specific charge e/m of electrons. K = cathode, A = anode.

Next we determine the mass m_e of the electron. This is not a direct measurement. In the various methods used, of which we discuss only one, the specific charge e/m_e, the charge per unit mass, of electrons is measured and then, using the experimentally measured value of e, m_e is calculated. As early as the last century e/m_e was determined by observing the electrostatic and magnetic deflection of fast electrons (cathode rays) in crossed electrostatic and magnetic fields. Figure 10 is a schematic illustration of a cathode-ray tube built by F. Braun, in which electrons originating at K pass through the electrostatic field between the condenser plates and simultaneously through the magnetic field between the external coils. The electrostatic and magnetic fields are mutually perpendicular. In the electric field the electrons are accelerated according to the equation

$$b = \frac{eE}{m_e} \tag{2-29}$$

where b is the acceleration. A time $t = l/v$ is required for electrons of velocity v to pass through the whole length l of the condenser. The electrons will be deflected through a distance

$$y = \frac{1}{2}bt^2 = \frac{eEl^2}{2m_e v^2} \tag{2-30}$$

perpendicular to the initial direction. The second equation necessary to determine the two unknown quantities v and e/m is obtained by observing the motion of the electrons in a magnetic field of strength H. Since the field is perpendicular to the initial direction of the electrons, they will describe a circle of radius

$$R = \frac{mcv}{eH} \tag{2-31}$$

The electrostatic deflection y and the angle through which the electron stream is deviated in passing through the magnetic field in a spiral path of radius R can be observed on the luminescent screen of the tube. The velocity v of the electrons can be eliminated from Eqs. (2-30) and (2-31), and e/m_e can be computed. The results of this and similar measurements are

$$\frac{e}{m_e} = (1.75936 \pm 0.00018) \times 10^8 \ \frac{\text{amp-sec}}{\text{gm}} \tag{2-32}$$

From (2-32) and (2-24) the mass of the electron is

$$m_e = (9.105 \pm 0.0001) \times 10^{-28} \ \text{gm} \tag{2-33}$$

From (2-13) and (2-33) we see that the mass of the electron is only 1/1,837 of the mass of the lightest atom, that of hydrogen. The small inertia of the electron resulting from its small mass is an important factor in many electronic instruments which will be considered later.

The mass of the electron given in (2-33) actually is its rest mass because the velocities in the experiment are small compared to the velocity of light. According to the theory of relativity the mass of an electron moving with a velocity v is given by

$$m_e = \frac{m_{e0}}{\sqrt{1 - (v^2/c^2)}} \tag{2-34}$$

The electrons emitted as β-particles in radioactive decay (page 275) reach almost 90 per cent of the velocity of light and from (2-34) it can be seen that their "dynamic mass" is more than twice their rest mass. We shall see in nuclear physics (Chap. 5) that electrons can be produced experimentally which have velocities more than 99 per cent of the velocity of light. The masses of these electrons must be calculated from the relativistic formula (2-34). They reach up to several hundred m_{e0}. It may be mentioned that formula (2-34) was experimentally confirmed early in this century by Kaufmann, Bucherer, and others, who determined e/m directly from fast β-rays.

In addition to its charge and mass the electron has, as Goudsmit and Uhlenbeck found in 1925, a constant mechanical angular momentum (spin) and an associated magnetic moment. This follows from the results of spectroscopic and magnetic experiments, discussed on page 154 and page 160. According to the classical picture an electron which rotates about its own axis with constant angular velocity has a magnetic moment equal to

$$\mu = \frac{eh}{4\pi m_e c} = 9.27 \times 10^{-21} \quad \text{gauss cm}^3 \tag{2-35}$$

caused by its rotating electric charge. This quantity is called the Bohr magneton. We shall meet this again when we consider atomic magnetism on page 152. This picture of the rotating electron loses some of its significance in modern quantum theory, as we shall see in Chap. 4. However, the angular momentum (spin) and magnetic moment (2-35) remain important characteristics of the electron.

Most recent experiments make the picture even more complicated since the magnetic moment is actually 0.13 per cent larger than the Bohr magneton, and it is interesting to note that within the experimental error this amount is equal to e^2/hc. This is usually explained by attributing to the electron, in addition to the moment associated with the rotating charge, an intrinsic magnetic moment.

In addition to the electronic properties discussed, we might ask about its diameter or radius, just as we did on page 18 for the atom. However, here we run into the same, or even greater, difficulties such as how to define and determine the radius of the electron. In collisions the electron acts only through its Coulomb field decreasing with $1/r^2$, so that no reasonable definition of the impact radius is possible. And we have no other experimental evidence of the structure or radius of the electron. Only in a rather artificial manner can a so-called "classical radius" of the electron be defined. If we compute from the mass-energy equivalence equation

$$E = m_e c^2 \tag{2-36}$$

the rest energy corresponding to the rest mass of the electron, and equate this to the electrostatic potential energy of a sphere having a charge e,

$$E_{\text{pot}} = \frac{e^2}{r} \tag{2-37}$$

we obtain the "classical radius" of the electron

$$r_e = \frac{e^2}{m_e c^2} = 2.82 \times 10^{-13} \quad \text{cm}\dagger \tag{2-38}$$

---

† In other calculations a factor 2 appears in the denominator resulting in the figure 1.41×10^{-13} cm. This, too, indicates a certain arbitrariness.

Two doubtful assumptions were made in this derivation. The first is the assumption that Coulomb's law is valid down to a distance of 10^{-13} cm, and the second, that the entire rest energy (2-36) of the electron consists of electrostatic energy. Neither assumption has been proved, nor are they very probable, so that the classical radius of the electron is of rather dubious significance. It seems certain, however, that the question of the eigenenergy of the electron and its meaning is very closely related to the present development of quantum theory and theory of elementary particles.

c. *Applications of Free Electrons. Electronic Instruments*

Now that we have learned all about the properties of normal electrons except for their behavior under extremely high impact energies which we shall discuss on page 326, we turn to a brief survey of the most important *applications* of free electrons. First of all we mention the use of electron rays as almost inertialess commutators. It is generally known that the fundamental element of amplifiers, radio receivers, and transmitters is the high vacuum radio tube (Fig. 11) which, in addition to a hot cathode and an anode, has one or more grids as control electrodes. A high negative potential on the grid forces the electrons back to the cathode, preventing a current flow to the anode. By varying the potential on the grid the flow of electrons can be controlled; the lower the grid potential, the more electrons pass and vice versa. The vacuum tube can be used as an amplifier, receiver, or transmitter depending on the circuit in which it is used. A vacuum tube with only a cathode and an anode can be used as a rectifier; since the cold anode cannot emit electrons, the current flows only when the cathode has a negative potential.

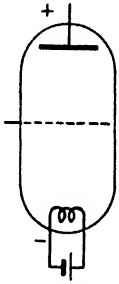


FIG. 11. Schematic representation of an electron tube with cathode, anode, and one grid.

Another instrument which makes use of the exceedingly small inertia of electrons has gained utmost importance for all kinds of electrical measurements, the cathode-ray oscilloscope. In this instrument, which is a development of Braun's tube (page 34), the cathode ray passes through two mutually perpendicular electrostatic fields (Fig. 12). The deviation of the electron beam can be observed on a fluorescent screen or photographed on a photographic plate perpendicular to the direction of the beam. If a periodic saw-toothed potential, causing a deviation proportional to time, is applied to one pair of plates, and an alternating potential which is to be investigated put on the other pair, a potential-versus-time curve (sine curve) will give traces on the screen from which the alternating potential can be studied in detail. Because electrons have almost zero inertia, extremely fast processes can be investigated in detail, for instance a spark discharge which occurs in 10^{-8} sec. The television tube is another variation of the Braun tube. In this tube the electron beam traverses the whole screen line by line 25 times a second and with varying intensity so that a picture can be received. We do not intend to go into the details of the physics of television in which a rapidly moving electron beam is also used in the transmitting tube (iconoscope). Among the many electronic instruments which have acquired considerable importance in science and industry, the electron microscope has become the most famous one. In this instrument, instead of the light beam in a light microscope, a beam of electrons from a hot cathode (Fig. 13) penetrates the object and is attenuated in varying degrees depending upon the absorption of the material which is used as the object. The divergent electron beam which leaves the object is then made convergent by electrostatic or magnetic fields (called electrostatic or magnetic lenses, Fig. 14) to form a picture of the object. Later on, a second lens combination magnifies this electron picture which is then projected on a fluorescent screen or photographed directly. Since the useful magnification of a light microscope is limited by its resolving power which is approximately equal to half a wavelength of the useful radiation, it cannot resolve, that is, separate, objects smaller than 0.2μ. As we shall see on page 197 the wavelength of electrons is dependent on their velocity; and with the potential (or electron velocity) used in the electron microscope objects smaller by a factor of 10^{-2} than those seen with visible light can be resolved. The electron microscope thus permits much finer detail, down to single large molecules, to be resolved. Its useful magnification therefore is at least 100 times greater than that of a light microscope.

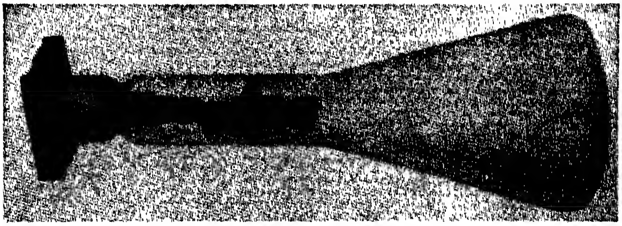


FIG. 12. Electron oscilloscope tube. At the right side of the system the two deflecting condensers, mutually perpendicular to each other, may be recognized.

In conclusion, we mention briefly that when fast electrons strike solid matter, such as a metal block, X-rays are produced, so that X-ray tubes are also electronic instruments. This short treatment will be sufficient to demonstrate how very important free electrons are for the whole field of physics and industry and that we can say with certainty that they will play an even more important role in the future.

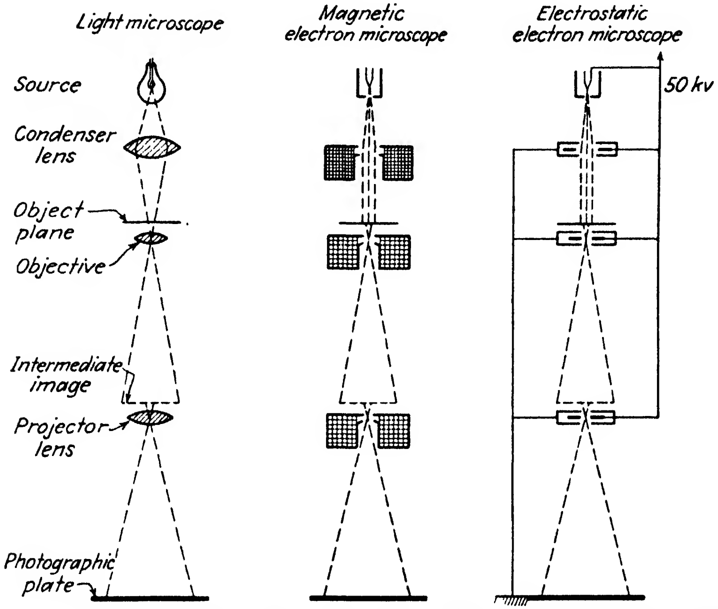


FIG. 13. Schematic sketches of light microscope, magnetic and electrostatic electron microscope.

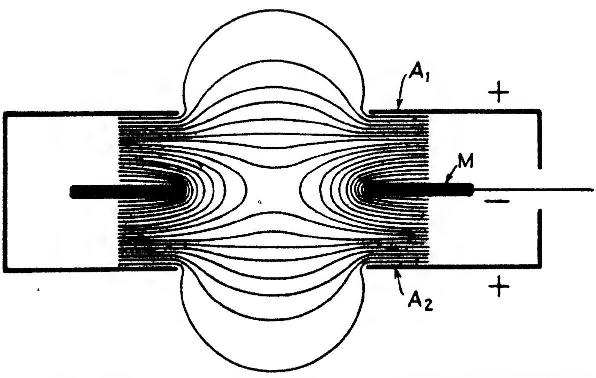


FIG. 14. Schematic sketch of the potential field of an electrostatic electron lens. A_1 and A_2 are external electrodes, M is a central electrode.

d. Free Ions

We now turn to free ions. We have already mentioned that positive ions are atoms or molecules which have lost one, and, in rare instances, two or more electrons from their shells. Negative ions are formed when an electron attaches itself to an electronegative atom (right side of the periodic table). Since the electron affinity (the binding energy between the electron and a neutral atom or molecule) is small, with exception of that of the halogen atoms, negative ions are not very important in gaseous processes. In the following we therefore will consider only positive ions. We shall designate ions by adding the appropriate number of + or − signs as superscripts to the symbol of the atom, thus speaking of a Ca^{++} ion.

The same methods used to produce electrons are available for producing positive ions. Electrons, and with them positive ions, are formed by impact ionization of fast atomic particles in the cloud chamber, ionization chamber, and Geiger counter (page 251). They are also produced by thermal ionization in superheated gases and vapors from which the ions can be withdrawn by a suitable arrangement of electric fields. The production of fast, directed ion beams in low-pressure gas discharges has gained special importance for mass spectroscopy. At sufficiently low pressures the ions formed by impact in the gas volume travel toward the cathode perpendicularly to its surface without being deviated by too frequent collisions with gas atoms. If the cathode is perforated according to Fig. 15, the accelerated ions pass through the cathode into an evacuated field-free chamber as so-called canal rays and can be studied there without being disturbed by electrons or other plasma particles. In high vacuum discharges of high potentials, for instance, in a condenser discharge through an evacuated capillary tube, highly charged ions can be produced, that is, ions with 3, 4, 5, and even up to 23 electrons removed. These have been called "stripped atoms" (page 116) because they have lost a considerable fraction of their original electron shell.

The production of positive ions at a solid surface by impact of fast electrons, as the counterpart of the production of electrons by impact of ions, occurs to a small extent at the anode in many discharges. In general this process has little significance. We shall deal with this case separately on page 475.

On the other hand there are two important methods of producing positive ions thermally. According to Kunsman certain mixtures of alkaline earths or alkaline compounds with other oxides emit metal ions when heated to about 900°C; for example $CsCO_3 + SiO_2 + Al_2O_3$

acts as a Cs^+ ion source. These Kunsman ion sources, whose yield depends greatly on the right "formation" by careful heating in certain gas atmospheres, are very practical for producing ions in a high vacuum. The second method of thermal production of positive ions applies only to the alkali and heavier alkaline-earth atoms. If these atoms fall on an incandescent platinum sheet, they are reflected as positive ions. We shall discuss this phenomenon more thoroughly on page 473.

In conclusion a special case of positive ions may be mentioned, the α-particles which are emitted by radioactive substances. These particles have been experimentally identified as doubly charged helium atoms, He^{++}. In nuclear physics, in addition to the α-particle, the hydrogen atomic ion is very important, for, like the α-particle, it is a pure nucleus and, when accelerated, it has a greater effect on other nuclei than other ions which have an electron shell and therefore a size comparable to that of normal atoms.

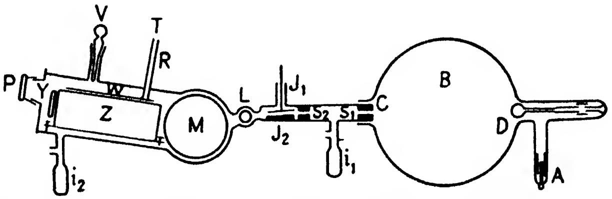


FIG. 15. Schematic representation of a canal-ray tube; C = cathode with hole, A = anode. At the left side an Aston mass spectrograph is connected to the canal-ray tube (see page 48 and Fig. 18).

We based our discussion of the ion on our knowledge of electrons and of the formation of electrons and positive ions by ionization of atoms, so that we can have no doubt about the nature of the positive ion. Historically, the direction of approach was reversed: positive charges of unknown nature were found in discharges, isolated as canal rays, and their nature determined by measuring their mass and charge. As the charge can be only one or a small integral multiple of the elementary quantum of electricity, the determination of the specific charge e/M was sufficient to determine M, where M is the mass of the ion. The measurement of e/M was made as with electrons (page 34) by measuring the deflection, in this case of canal rays, in crossed electrostatic and magnetic fields. By means of such measurements, especially by W. Wien, it was proved beyond any doubt that free ions from gas discharges or any other ion sources were nothing else but charged atoms or molecules; this discovery served also as proof for the existence of individual atoms. In the second section following this one we shall discuss the measurement of the masses of isotopes, the so-called mass spectroscopy, by the above methods.

2-5. Brief Survey of the Structure of the Atomic Nucleus

In this section we present a brief survey of the structure of the atomic nucleus, anticipating results from a more thorough discussion in Chap. 5. We need this information in order to understand the phenomena of isotopes which will be discussed in the following section.

We have already learned from the scattering experiments, page 22, that the diameter of the nucleus is approximately 10^{-12} cm, or one ten-thousandth of that of an atom, and that almost all of the mass of the atom is concentrated in the nucleus. From the nuclear volume and the absolute mass of the atom, page 18, one can thus easily compute that the nucleus has an inconceivably large density, of the order of 10^{14} gm cm^{-3}. One cubic centimeter of nuclear materials weighs approximately 100 million tons! It is not improbable that in the interior of certain giant fixed stars there exist large masses of matter of comparable density.

Each atomic nucleus is characterized by its atomic weight, its mass number (the unit being roughly 1/16 of the mass of the oxygen atom), and the atomic number which indicates the number of positive charges the nucleus carries. We have already mentioned Prout's hypothesis that all elements were built up of hydrogen atoms. Up to 1932 it was the general belief that a nucleus of mass number A was made up of A nuclei of hydrogen, called protons. Since the number of positive charges Z on the nucleus was about half as large as the mass number A, the nucleus had to be regarded as having $A - Z$ electrons, the so-called nuclear electrons, in the nucleus, in order to arrive at the correct nuclear charge Z. We shall consider in detail on page 270 the evidence against the existence of electrons in the nucleus. The discovery of neutrons (uncharged nuclear particles of mass and size approximately equal to the proton), which we will discuss in detail on page 289, made the hypothesis of nuclear electrons avoidable. Today we know for certain that the nucleus consists of Z protons and $A - Z$ neutrons which are bound together by exchange forces of a type which was first explained by quantum mechanics (page 228). We also know that, because of a saturation effect of these nuclear binding forces, two protons and two neutrons form an α-particle; and the emission of such α-particles by radioactive elements (nuclei) confirms this theoretical conclusion. We shall learn all the details about the structure of the atomic nucleus in Chap. 5.

2-6. Isotopes

a. The Discovery of Isotopes and Their Significance for the Atomic Weight Problem

In discussing atomic weights (page 14) we have shown that neither Prout's hypothesis of H atoms as the constituents of all atoms nor the new concept of the nucleus as consisting of protons and neutrons is compatible with atomic weights which are nonintegral numbers of 1/16 oxygen mass. However, there are a number of such elements in the periodic table (page 16) whose atomic weights are not integral numbers, for example, chlorine with an atomic weight of 35.46. This difficulty was resolved, as we briefly mentioned early in the chapter, when it was discovered that atoms with the same nuclear charge and, therefore, of the same element, could have different atomic weights!

Isotopy was discovered after 1907 when experiments with radioactive elements and their disintegration products uncovered many elements which had different decay periods and disintegration products, but which could not be separated chemically. Thorium, radiothorium, and ionium, all having the atomic number 90, were the first three atoms discovered which differed in their physical properties but were chemically identical. They were called isotopes. The first evidence that elements of nonradioactive type had isotopes was uncovered when Hönigschmid in 1914 measured the atomic weight of a sample of lead which was an end product of the uranium ores and found it to be 206.05, whereas the atomic weight of the lead derived from theorium ores was 207.90. We shall see (page 274 and Fig. 137) that the lead isotope derived from the radioactive uranium family actually has an atomic weight of 206 while the isotope derived from the thorium family has an atomic weight of 208. At the same time Thomson, during his study of neon canal rays in crossed electrostatic and magnetic fields, discovered that neon did not consist of atoms of uniform mass but had three isotopes. In the following years Aston, using his famous first mass spectrograph, proved that isotopes are not exceptions but rather that most elements were actually mixtures of isotopes.

With these discoveries the question of atomic weights appeared in an entirely different light. The atomic weight determined chemically is no longer a constant of the atom. It is the average value of the atomic weights and relative abundances of the isotopes of the elements. This also holds true for the element used as a reference for atomic masses, oxygen, which is now known to be a mixture of isotopes of atomic weights 16, 17, and 18. It is only a fortunate coincidence that the isotope of mass 16, O^{16}, has a relative abundance of 99.76 per cent as compared to 0.04 and 0.20 per cent for the other isotopes. The chemical scale of atomic weights, taking the average mass of the isotopic mixture of oxygen as 16.0000, and the physical scale, taking the mass of the isotope $O^{16} = 16.0000$, thus differ by less than three parts in ten thousand. To be precise, the atomic weight in the chemical scale is 0.0278 per cent smaller than that in the physical scale. It is important to note that, in general, the relative abundance of the isotopes of the elements is constant on our earth and in the whole universe so that, even if the chemical atomic weight is no longer a constant of the atom, it is the *average atomic weight of the natural mixture of isotopes* and is a characteristic property of any particular element. The only exceptions to the above rule are the end products of the radioactive elements which result from nuclear transformations. The outstanding example is lead, the end product of three radioactive series, which has an average atomic weight that varies according to the locality where it is found.

b. *The Explanation of Isotopes and Their Properties*

From what is known about the structure of the nucleus, there can be no doubt about the explanation of isotopes: the isotopes of an element have the same number of protons but a different number of neutrons in the nucleus. The most famous and, scientifically, one of the most important isotopes is that of the hydrogen atom. Its nucleus consists of one proton and one neutron and thus has the mass number 2, whereas the usual nucleus with charge and mass 1 is just a proton. We shall see in Chap. 5, when we take up the systematic treatment of nuclear structure, which factors determine whether an element has many or a few or only one stable isotope. Quite generally there exists for every element one definite proportion of protons to neutrons for which stability is greatest, whereas too great an excess or deficit of neutrons causes instability of a nucleus. We shall see when we discuss artificial radioactivity that such nuclei revert to a stable state by ejecting an electron or a positron. Unstable isotopes can be produced by artificial nuclear transformations, so that in addition to the stable isotopes, every element has a number of unstable, radioactive isotopes. Oxygen, for instance, has, besides the stable isotopes O^{16}, O^{17}, and O^{18} the unstable, radioactive isotopes O^{15} and O^{19}. We shall consider this subject in detail in Chap. 5.

In order to understand completely the behavior of any chemical element it is no longer sufficient to know just the chemical atomic weight, it is also necessary to know the mass and relative abundance of all its stable isotopes. Later on we shall take up in detail the methods for determining their mass and relative abundance. At this time we shall only mention one very important fact: that all isotopes, relative to $O^{16} = 16.0000$, have atomic weights which are integral numbers within 1 per cent. This important result is, as was pointed out on page 14, in complete agreement with the theory that the nucleus is built up of protons and neutrons.

Now that we know that the isotopes of an element differ from one another only in the number of neutrons in their nuclei, we can draw several important conclusions about the general properties of isotopes. In the first place it is evident that a variation in the number of neutrons in the nucleus, because of the neutron's electric neutrality, has no influence on the binding forces between the positive nucleus and the negative electron shells, so that the electron shells and the chemical and physical properties which depend on them alone are the same for all isotopes of an element. Therefore, we can expect a different behavior of isotopes only in those properties which depend on their mass. These properties then can be used to distinguish between the isotopes of the same element. This difference in properties is the more pronounced, the greater the difference in the masses of the isotopes of a particular element; isotopes of the light nuclei thus will be easiest to separate. Hence the importance of the heavy hydrogen isotope, ${}_1H^2$ (the number of charges is indicated in the subscript and the mass number in the superscript), becomes clear. Because of this special importance, the atom of heavy hydrogen has been given a separate name, deuterium, with the symbol D. The nucleus of ${}_1H^2 = {}_1D^2$ is called a deuteron and is designated by the symbol d. The mass of deuterium, ${}_1D^2 = {}_1H^2$, is twice that of the normal hydrogen atom, ${}_1H^1$, while the greatest difference in the masses of oxygen isotopes, $O^{18} - O^{16}$, is only 12 per cent, and that of uranium, $U^{238} - U^{235}$, is only 1 per cent. Similarly the specific charges e/M of the positive ions of the isotopes must vary by corresponding amounts. The measurement of e/M therefore provides an exact method for determining the masses of the isotopes. Among the properties of the atom which depend upon the mass are the evaporation velocity and the coefficient of diffusion, which can be used to distinguish between isotopes and to separate them. We shall investigate the small but significant differences which different masses of isotopes cause in atomic and molecular spectra in Chaps. 4 and 5. The difference in the masses of H and D is the reason why there is a substantial difference in the moments of inertia as well as in the zero-point energy of the molecules H_2 and HD, and why there is a difference in the behavior of the hydrogen compounds NH_3 and ND_3, and CH_4 and CD_4. It also explains the difference in the heats of vaporization, molecular heats, molar volumes, chemical constants, dissociation energies, and vibration energies of these molecules. As an example, the heats of vaporization of H_2 and D_2 are 220 and 302 cal, respectively, but the corresponding difference for H_2O and D_2O is only 3½ per cent.

c. The Determination of the Masses and Relative Abundances of Isotopes

To determine the mass and relative abundance of isotopes we may use the methods of optical spectroscopy. The difference in mass can be determined from the difference in the wavelengths of the spectral lines belonging to the various isotopes, and the relative abundance by measuring the relative intensities of these lines. Investigation of isotopes by spectroscopy, in particular by band spectroscopy, has contributed much to our present knowledge. We shall consider these optical methods of isotope investigation in detail on pages 181 and 398.

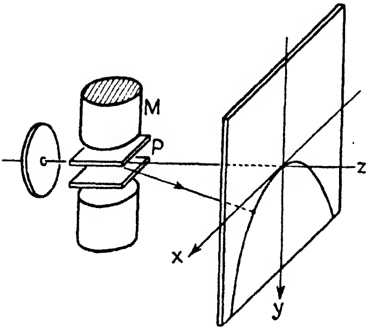


FIG. 16. Schematic representation of Thomson's parabola method for determining ion masses. P = condenser plates, M = magnetic pole pieces. (*After Brüche and Recknagel.*)

However, the classical method for investigating isotopes, introduced by Thomson and Aston and perfected by Mattauch, is the method of determining e/M, the so-called mass spectroscopy.

The parabolic method used by Thomson in 1913 is the oldest method used which gave reliable results. Instead of the usual crossed fields, Thomson used superimposed electrostatic and magnetic fields (Fig. 16) with the condenser plates between the pole faces. Let the direction of the canal rays be the z axis and the direction of the field be the y axis, then from (2-30) the deviation resulting from the electrostatic field acting over a distance l is

$$Y = \frac{eEl^2}{2Mv^2} \tag{2-39}$$

The deviation due to the magnetic field is perpendicular to y and z, that is, in the x-direction. The radius of the orbit described by the particles is, from (2-31),

$$R = \frac{Mcv}{eH} \tag{2-40}$$

or

$$X = \frac{1}{2}\frac{dv}{dt}t^2 = \frac{1}{2}\left(\frac{v^2}{R}\right)\left(\frac{l}{v}\right)^2 = \frac{l^2}{2R} = \frac{eHl^2}{2Mcv} \tag{2-41}$$

Since Y is proportional to $1/v^2$ and X is proportional to $1/v$, the trace of all particles of the same specific charge e/M on the phosphorescent screen or photographic plate is a parabola,

$$X^2 = 2pY = \frac{e}{2M}\frac{H^2l^2}{Ec^2}Y \tag{2-42}$$

The slope of the curve is used to determine the value of e/M and from it the mass of a particular ion. The intensities of the individual parabolas, Fig. 17, correspond to the relative abundance of the various ions in the mixture. The power of resolution for ions of various masses depends on the intensity of the canal rays which determine the smallest applicable diameter of the ion beam.

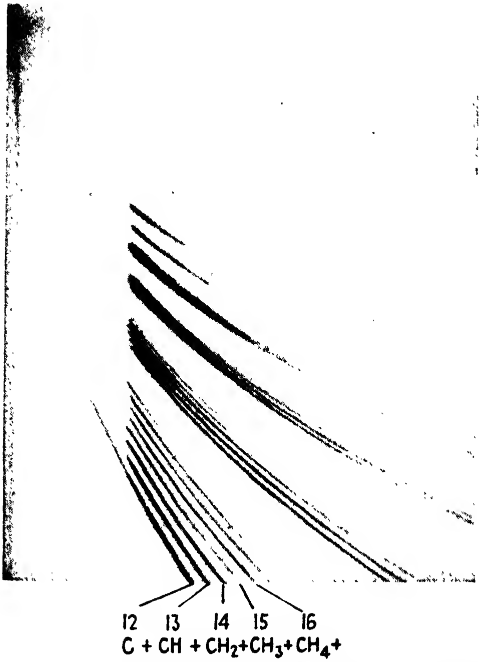


FIG. 17. Mass-spectroscopic analysis of a hydrocarbon mixture by the Thomson parabola method. (*Photograph by Conrad.*)

The second mass-spectroscopic method for determining e/M, and from it the mass of isotopes, uses crossed electrostatic and magnetic fields. It has been developed into a precision method. Figure 18 shows the first mass spectrograph built by Aston in 1917; Fig. 19 is a schematic diagram of the apparatus. It corresponds to an optical spectrograph in that it has an entrance slit, separates the rays in an electrostatic and magnetic field, and finally collects the beams of different mass at different points on a photographic plate or receiver. The canal rays pass through a fine hole S_1, bored through the cathode of a discharge tube, and through a second collimating hole, S_2. The rays then pass through the electric field E and are deflected through an angle α. After passing through a comparatively wider slit, K_2, they traverse the magnetic field H wherein they are deflected in an opposite direction through an angle β, after which they impinge on a photographic plate W. Aston realized that sufficient intensity for an exposure could be obtained only if slit K_2 were not too narrow, and that sharp lines, and thus a good separation of neighboring masses, would be possible only if by the focusing action of the magnet fields all ions of the same mass, regardless of their velocity, arrived at the same place on the photographic plate. To accomplish this "velocity focusing" there must be a definite relation between the angles α and β, as well as between the distances between the centers of the electrostatic and magnetic fields, and between these and the image points on the plate. By using the expressions for the deflections (2-39) and (2-41), which now are both in the y-direction, since the fields are crossed, and substituting

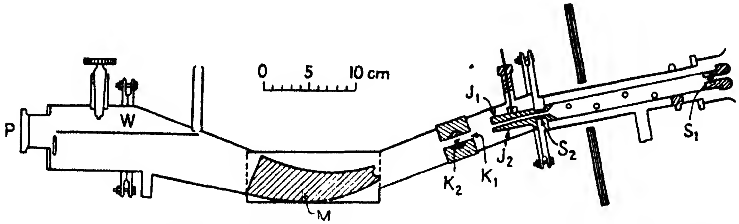


FIG. 18. Improved mass spectrograph of Aston. S_1, S_2, K_1, and K_2 = diaphragms, J_1 and J_2 = deviating condenser plates, M = magnetic pole pieces, W = photographic plate.

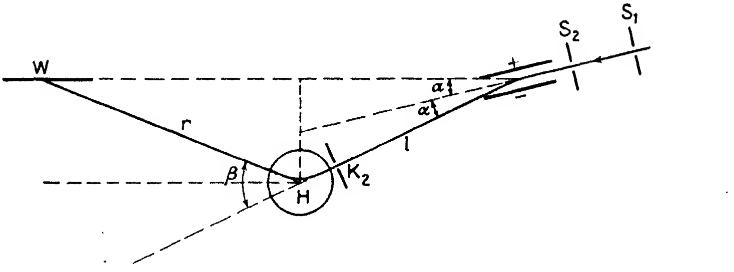


FIG. 19. Schematic diagram of an ion beam in Aston's mass spectrograph. S_1, S_2, and K_2 = slits, H = magnetic field perpendicular to the plane of the paper, α = deflection in the electric field, β = deflection in the magnetic field, W = photographic plate.

for the velocity v the expression which includes the accelerating potential V,

$$eV = \frac{M}{2}v^2 \qquad v = \sqrt{\frac{2eV}{M}} \tag{2-43}$$

we can obtain for the deflection angles α and β of the ion rays

$$\alpha = \frac{dY_1}{dx} = \frac{eEa}{Mv^2} = \frac{Ea}{2V} \tag{2-44}$$

and

$$\beta = \frac{dY_2}{dx} = \frac{eHb}{Mvc} = \frac{Hb\sqrt{e}}{c\sqrt{2MV}} \tag{2-45}$$

where a and b are the effective diameters of the electrostatic and magnetic fields, respectively. If there is a variation in the accelerating voltage which has a definite magnitude dV at K_2, the resulting variation in the angles α and β is given by differentiating (2-44) and (2-45) with respect to V.

$$d\alpha = -\frac{dV}{V}\alpha \tag{2-46}$$

and

$$d\beta = -\frac{dV}{2V}\beta \tag{2-47}$$

According to Fig. 19 the total deflection of the ion rays, measured on a screen perpendicular to the initial ray direction, is

$$Y_1 \pm Y_2 = Y = \alpha(l + r) - \beta r \tag{2-48}$$

and the broadening of each mass line, resulting from the variation in velocity dV

$$dY = \left[\alpha(l + r) - \frac{\beta r}{2}\right]\frac{dV}{V} \tag{2-49}$$

Evidently this broadening must vanish when the expression in the brackets vanishes, i.e., for

$$r = \frac{2\alpha l}{\beta - 2\alpha} \tag{2-50}$$

Equation (2-50) is thus the condition for velocity focusing. If it is fulfilled, the mass spectrograph gives sharp images of all ions of the same mass at a point on the photographic plate W.

Figure 20 shows several photographs by Aston with mass numbers which permit us to judge the selectivity of his instrument. These photographs have led to the discovery of a large part of the now-known isotopes.

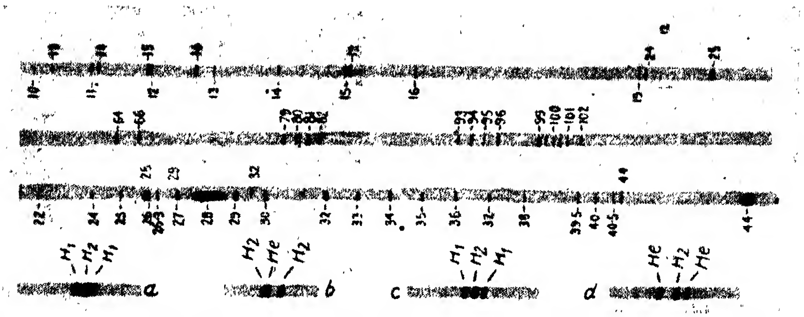


FIG. 20. Mass spectra of Aston which were used to determine masses and relative abundances of isotopes.

Aston's mass spectrograph has been modified in various ways. Dempster showed that the magnetic field can also be used to focus ions which have different initial directions on one point of the plate (direction focusing). The final improvement was accomplished by Mattauch in 1934, when, by using carefully computed distances and angles of deflection, he succeeded in simultaneously focusing ions of different initial velocities as well as directions on a plane photographic plate (Fig. 21). Slit K, which is a perforation of the cathode, can be moderately wide because the slit S which is imaged on the plate regulates the sharpness of the lines. The ion beam is first spread out in a radial electrostatic field of 31°50′ with respect to velocities and then, after passing through a reasonably wide diaphragm B, enters the magnetic field where it is deflected through an angle of 90°. By the action of both fields, the initial beam is now spread out separated into the various masses, and these are focused with respect to both velocity and direction. Figure 22 presents an example of a mass spectrum which can

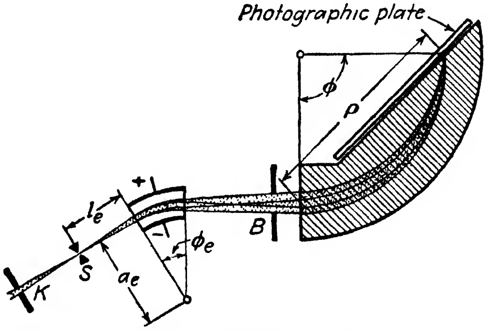


FIG. 21. Improved mass spectrograph of Mattauch with velocity and direction focusing. ϕ_e = deviation in the electric field, ϕ = deviation in the magnetic field.

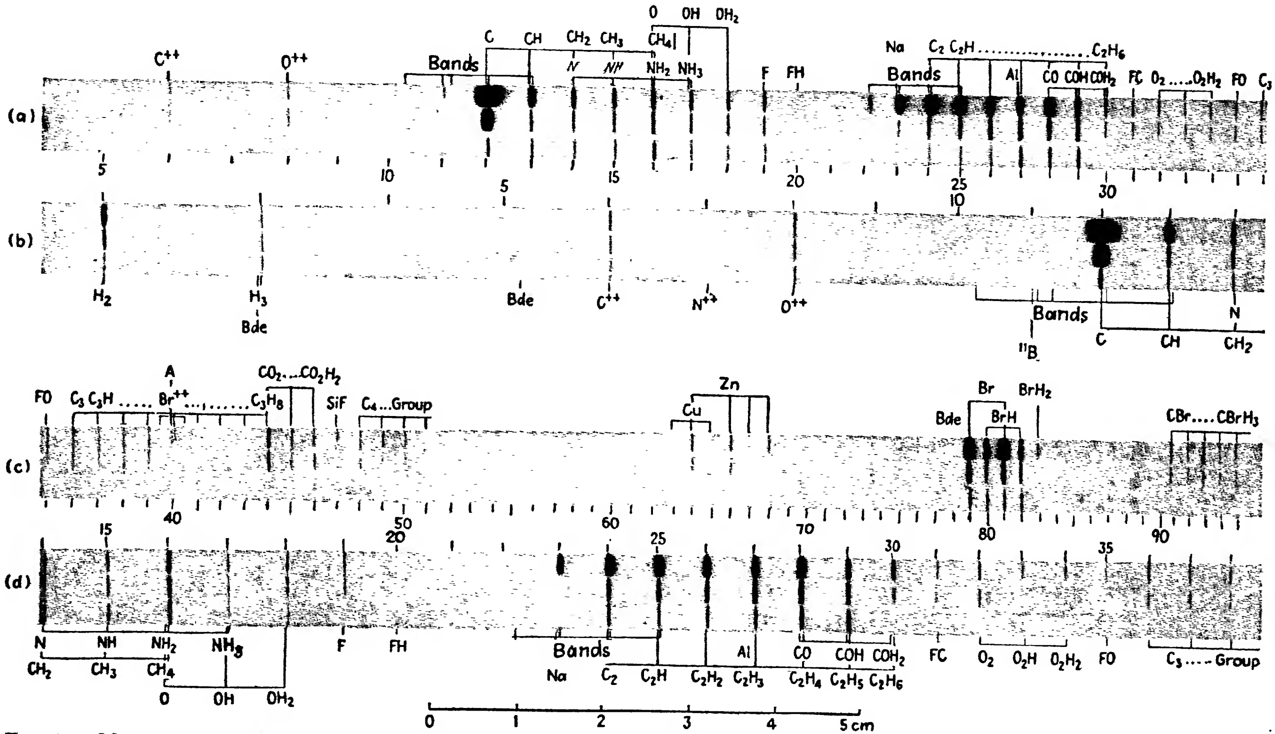


FIG. 22. Mass spectra of Mattauch. Note the greatly improved definition of the lines (compared with Fig. 20), which allows more accurate mass determination.

compete with any optical spectrum as to the definition of its lines. Figure 23 is an especially fine example of the high resolution possible with this instrument. With this construction the distance between different lines is proportional to the square roots of the mass differ-

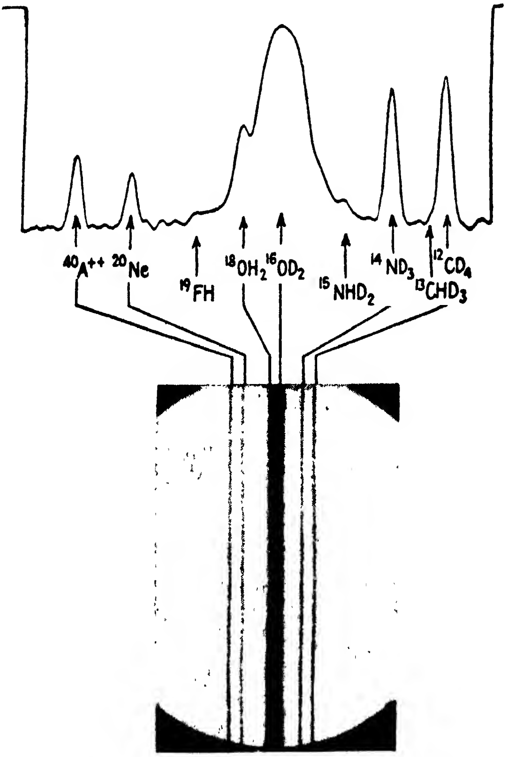


Fig. 23. Fine-structure mass spectrum of Mattauch with densitometer record, showing the definite separation, thus allowing the measurement of nine different ions of mass number 20 whose masses range from 19.9878 to 20.0628 mass units.

ences. For calibrating the mass spectrograph, mass lines of CO^+ and ions of some simple hydrocarbons are used, which are usually present as minute impurities. Ion masses, and thus atomic weights, can be measured in the mass spectrograph with an accuracy within 10^{-4} mass units. We cannot realize the great importance of these measurements before we consider the mass defects of the nuclei and their bearing on nuclear stability (page 271). The relative abundance of the isotopes, noted in column 5 of Table 3, can be obtained by measuring the intensities of the lines, e.g., those of the Zn isotopes in Fig. 22.

Table 3. The Stable Isotopes of the Elements (Including the Most Stable Isotopes of Those Elements Which Do Not Have Stable Ones)

Z	Element	Mass number, A	Neutron number, A − Z	Relative abundance, per cent	Isotope mass, O = 16.0000	Atomic weight, chemical scale (1940)
0	*n*	1	1		1.00895	
1	*H*	1	0	99.986	1.00813	1.0080
	(D)	2	1	0.014	2.01472	
2	He	3	1	10^{-6}	3.01700	4.003
		4	2	100	4.00386	
3	Li	6	3	7.9	6.01692	6.940
		7	4	92.1	7.01816	
4	Be	9	5	100	9.01496	9.02
5	B	10	5	19.8	10.01617	10.82
		11	6	80.2	11.01290	
6	C	12	6	98.9	12.00388	12.010
		13	7	1.1	13.00756	
7	N	14	7	99.62	14.00753	14.008
		15	8	0.38	15.00487	
8	O	16	8	99.76	16.00000	16.0000
		17	9	0.04	17.00450	
		18	10	0.20	18.00485	
9	F	19	10	100	19.00447	19.00
10	Ne	20	10	90.00	19.99890	20.183
		21	11	0.27	20.99980	
		22	12	9.73	21.99862	
11	Na	23	12	100	22.99644	22.997
12	Mg	24	12	77.4	23.9930	24.32
		25	13	11.5	24.9946	
		26	14	11.1	25.9901	
13	Al	27	14	100	26.9907	26.97
14	Si	28	14	89.6	27.9870	28.06
		29	15	6.2	28.9865	
		30	16	4.2	29.9839	
15	P	31	16	100	30.9844	30.98
16	S	32	16	95.04	31.9825	32.06
		33	17	0.74	32.9819	
		34	18	4.2	33.9798	
		36	20	0.016	?	
17	Cl	35	18	75.4	34.97884	35.457
		37	20	24.6	36.9777	
18	A	36	18	0.31	35.9973	39.944
		38	20	0.06	37.9746	
		40	22	99.63	39.9755	
19	K	39	20	93.44	38.976	39.096
		40	21	0.012	?	
		41	22	6.55	?	

Table 3. The Stable Isotopes of the Elements (Continued)

Z	Element	Mass number, A	Neutron number, $A - Z$	Relative abundance, per cent	Isotope mass, 0 = 16.0000	Atomic weight, chemical scale (1940)
20	Ca	40	20	96.96		40.08
		42	22	0.64		
		43	23	0.15		
		44	24	2.07		
		46	26	0.003		
		48	28	0.185		
21	Sc	45	24	100	44.96977	45.10
22	Ti	46	24	7.95		47.90
		47	25	7.75		
		48	26	73.45	47.966	
		49	27	5.51	48.964	
		50	28	5.34	49.963	
23	V	51	28	100		50.95
24	Cr	50	26	4.49		52.01
		52	28	83.78	51.959	
		53	29	9.43		
		54	30	2.30		
25	Mn	55	30	100		54.93
26	Fe	54	28	5.84	53.959	55.85
		56	30	91.68	55.957	
		57	31	2.17		
		58	32	0.31		
27	Co	59	32	100		58.94
28	Ni	58	30	67.4	57.9597	58.69
		60	32	26.7	59.949	
		61	33	1.2	60.9540	
		62	34	3.8	61.9496	
		64	36	0.88	63.9474	
29	Cu	63	34	68	62.957	63.57
		65	36	32	64.955	
30	Zn	64	34	50.9	63.957	65.38
		66	36	27.3	65.953	
		67	37	3.9		
		68	38	17.4	67.955	
		70	40	0.5	69.954	
31	Ga	69	38	61.5	68.956	69.72
		71	40	38.5	70.954	
32	Ge	70	38	21.2		72.60
		72	40	27.3		
		73	41	7.9		
		74	42	37.1		
		76	44	6.5		
33	As	75	42	100		74.91

Table 3. The Stable Isotopes of the Elements (Continued)

Z	Element	Mass number, A	Neutron number, $A - Z$	Relative abundance, per cent	Isotope mass, $0 = 16.0000$	Atomic weight, chemical scale (1940)
34	Se	74	40	0.9		78.96
		76	42	9.5		
		77	43	8.3		
		78	44	24.0		
		80	46	48.0		
		82	48	9.3		
35	Br	79	44	50.6		79.916
		81	46	49.4		
36	Kr	78	42	0.35	77.945	83.7
		80	44	2.01		
		82	46	11.52	81.938	
		83	47	11.52		
		84	48	57.13	83.939	
		86	50	17.47	85.939	
37	Rb	85	48	72.8		85.48
		87	50	27.2		
38	Sr	84	46	0.56		87.63
		86	48	9.86		
		87	49	7.02		
		88	50	82.56		
39	Y	89	50	100		88.92
40	Zr	90	50	48		91.22
		91	51	11.5		
		92	52	22		
		94	54	17		
		96	56	1.5		
41	Cb	93	52	100		92.91
42	Mo	92	50	14.9		95.95
		94	52	9.4	93.945	
		95	53	16.1	94.945	
		96	54	16.6	95.946	
		97	55	9.65	96.945	
		98	56	24.1	97.944	
		100	58	9.25	99.939	
43	Tc	99	56			
44	Ru	96	52	(5)	95.945	101.7
		98	54			
		99	55	(12)	98.944	
		100	56	(14)		
		101	57	(22)		
		102	58	(30)		
		104	60	(17)		
45	Rh	103	58	100	102.949	102.91

Table 3. The Stable Isotopes of the Elements (Continued)

Z	Element	Mass number, A	Neutron number, $A - Z$	Relative abundance, per cent	Isotope mass, $0 = 16.0000$	Atomic weight, chemical scale (1940)
46	Pd	102	56	0.8		106.7
		104	58	9.3		
		105	59	22.6		
		106	60	27.2	105.946	
		108	62	26.8		
		110	64	13.5	109.944	
47	Ag	107	60	52.5	106.950	107.880
		109	62	47.5	108.949	
48	Cd	106	58	1.4		112.41
		108	60	1.0		
		110	62	12.8		
		111	63	13.0		
		112	64	24.2		
		113	65	12.3		
		114	66	28.0		
		116	68	7.3		
49	In	113	64	4.5		114.76
		115	66	95.5		
50	Sn	112	62	1.1		118.70
		114	64	1.8		
		115	65	0.4		
		116	66	15.5	115.943	
		117	67	19.1		
		118	68	22.5	117.940	
		119	69	9.8	118.938	
		120	70	28.5		
		122	72	5.5	121.946	
		124	74	6.8	123.945	
51	Sb	121	70	56		121.76
		123	72	44		
52	Te	122	70	2.9		127.61
		123	71	1.6		
		124	72	4.5		
		125	73	6.0		
		126	74	19.0		
		128	76	32.8		
		130	78	33.1		
53	I	127	74	100		126.92
54	Xe	124	70	0.094		131.3
		126	72	0.088		
		128	74	1.91		
		129	75	26.23	128.946	
		130	76	4.06		

Table 3. The Stable Isotopes of the Elements (Continued)

Z	Element	Mass number, *A*	Neutron number, *A* − *Z*	Relative abundance, per cent	Isotope mass, 0 = 16.0000	Atomic weight, chemical scale (1940)
		131	77	21.18		131.3
		132	78	26.98	131.946	
		134	80	10.55		
		136	82	8.95		
55	Cs	133	78	100		132.91
56	Ba	130	74	0.101		137.36
		132	76	0.097		
		134	78	2.42		
		135	79	6.6		
		136	80	7.8		
		137	81	11.3		
		138	82	71.7		
57	La	139	82	100	138.955	138.92
58	Ce	136	78	Rare		140.13
		138	80	Rare		
		140	82	89		
		142	84	11		
59	Pr	141	82	100		140.92
60	Nd	142	82	25.95		144.27
		143	83	13.0		
		144	84	22.6		
		145	85	9.2		
		146	86	16.5	145.964	
		148	88	6.8	147.964	
		150	90	5.95	149.970	
61	Pm	147	86			
62	Sm	144	82	3		150.43
		147	85	17		
		148	86	14		
		149	87	15		
		150	88	5		
		152	90	26		
		154	92	20		
63	Eu	151	88	49		152.0
		153	90	51		
64	Gd	152	88	0.2		156.9
		154	90	1.5		
		155	91	21	154.977	
		156	92	22	155.977	
		157	93	17	156.976	
		158	94	22	157.976	
		160	96	16	159.976	
65	Tb	159	94	100		159.2

Table 3. The Stable Isotopes of the Elements (Continued)

Z	Element	Mass number, A	Neutron number, $A - Z$	Relative abundance, per cent	Isotope mass, $O = 16.0000$	Atomic weight, chemical scale (1940)
66	Dy	158	92	0.1		162.46
		160	94	1.5		
		161	95	22		
		162	96	24		
		163	97	24		
		164	98	28		
67	Ho	165	98	100		164.94
68	Er	162	94	0.25		167.2
		164	96	2		
		166	98	(35)		
		167	99	(24)		
		168	100	(29)		
		170	102	(10)		
69	Tm	169	100	100		169.4
70	Yb	168	98	0.06		173.04
		170	100	4.21		
		171	101	14.26		
		172	102	21.49		
		173	103	17.02		
		174	104	29.58		
		176	106	13.38		
71	Lu	175	104	97.5		174.99
		176	105	2.5		
72	Hf	174	102	(0.3)		178.6
		176	104	(5)		
		177	105	(19)		
		178	106	(28)		
		179	107	(18)		
		180	108	(30)		
73	Ta	181	108	100		180.88
74	W	180	106	0.2		183.92
		182	108	22.6		
		183	109	17.3		
		184	110	30.1		
		186	112	29.8		
75	Re	185	110	38.2		186.31
		187	112	61.8		
76	Os	184	108	0.018		190.2
		186	110	1.59		
		187	111	1.64		
		188	112	13.3		
		189	113	16.2		
		190	114	26.4	190.038	
		192	116	40.9	192.038	

Table 3. The Stable Isotopes of the Elements (Concluded)

Z	Element	Mass number, A	Neutron number, $A - Z$	Relative abundance, per cent	Isotope mass, $0 = 16.0000$	Atomic weight, chemical scale (1940)
77	Ir	191	114	38.5	191.038	193.1
		193	116	61.5	193.039	
78	Pt	192	114	0.8		195.23
		194	116	30.2	194.039	
		195	117	35.3	195.039	
		196	118	26.6	196.039	
		198	120	7.2	198.044	
79	Au	197	118	100	197.039	197.2
80	Hg	196	116	0.15		200.61
		198	118	10.12		
		199	119	17.04		
		200	120	23.25		
		201	121	13.18		
		202	122	29.54		
		204	124	6.72		
81	Tl	203	122	29.1	203.059	204.39
		205	124	70.9	205.059	
82	Pb	204	122	1.5	204.058	207.21
		206	124	23.6		
		207	125	22.6		
		208	126	52.3	208.057	
83	Bi	209	126	100	209.055	209.00
84	Po	210	126			
85	At	211	126			
86	Rn	222	136			222
87	Fr	223	136			
88	Ra	226	138			226.05
89	Ac	227	138			
90	Th	232	142	100		232.12
91	Pa	231	140			231
92	U	235	143	0.72		238.07
		238	146	99.28		
93	Np	237	144			
94	Pu	239	145			
95	Am	241	146			
96	Cm	242	146			

d. Methods of Separating Isotopes

Originally, isotopes were studied to increase our knowledge of atomic structure; today the isotopes are a great aid to molecular physicists, chemists, and biologists, for an isotope can be used to identify and trace one individual atom among a great number of chemically similar ones. Therefore, science and technology are greatly interested in obtaining pure isotopes, i.e., in separating as completely as possible the various isotopes of an element. For many purposes, or as a first step toward complete separation, it is sufficient to obtain a strong concentration of a rare isotope.

The most exact method of separation, though suited only for small quantities, is that of mass spectroscopy. A mass spectrograph with large, intensively emitting ion sources is used. Various masses are separated into beams and then concentrated by suitable electrostatic and magnetic lenses at different places where the separated isotopes are collected. In this manner about 10^{-6} gm, and in some favorable cases 10^{-4} gm, of absolutely pure isotopes can be produced per hour.

If larger quantities are required, they have to be obtained by enriching the concentration rather than by complete separation of the desired isotope. One method utilizes the difference in vapor pressure of isotopes, especially of those of the light elements with a large relative mass difference. By fractional distillation at the lowest possible temperature a considerable change in the relative abundances of an isotope mixture can be obtained. However, it is not possible to use this method for an actual separation. A second method was developed by G. Hertz. In this method the concentration of a component of a mixture of gaseous isotopes is increased by successive diffusion through clay tubes or by passing the mixture through a series of diffusion pumps. In this manner, for instance, the concentration of A^{38} can be increased from 0.06 per cent to about 50 per cent. The method can be applied to all gases and vapors. The method used in the United States to increase the concentration of heavy hydrogen in water, besides fractional distillation, depends on the catalytic exchange of heavy hydrogen between hydrogen gas and water vapor. By properly applying a catalyst, which is still an undisclosed secret, the equilibrium density of deuterium in the water becomes three times as great as it is in the gas. If countercurrents of water vapor and hydrogen gas are allowed to bubble through the catalyst, and if the process is repeated in numerous successive stages, a high concentration of deuterium will be found in the condensed water. The production of pure deuterium, which is required for many applications today, is carried out exclusively by electrolysis of water. Because of a number of interdependent reactions which depend upon mass (variation of ion mobility, the facility with which ions give up their charge to the electrodes, etc.) the light H_2O molecule, predominantly, is decomposed, leaving the HDO and D_2O molecules in the water. By applying a number of empirically proved rules concerning the proper choice of the electrolyte, electrodes, etc., it is possible to obtain from 100 qt of water, by expending 1,000 kwh of electrical energy, 1 gm of D_2O of 98 per cent purity. For laboratory purposes, D_2O now can be purchased and D_2 obtained from it by electrolytic decomposition. Generally, when heavy hydrogen is concentrated in water the abundance of the heavy oxygen isotopes O^{17} and O^{18} is also increased. For elements with a smaller relative mass difference of isotopes the electrolytic method of separation is not very satisfactory.

The present standard method for producing large quantities of many types of isotopes was first suggested several years ago by Clusius. This successful method, the separation-tube process, is based on the principle of thermal diffusion, which was known but not appreciated for a long time: If a temperature gradient is maintained in a gas mixture the lighter atoms or molecules diffuse in the direction of the higher temperature, the heavier atoms toward the region of lower temperature. Since, in equilibrium, ordinary diffusion tends to compensate this effect, Clusius and Dickel made use of an artifice to develop a practical method of separation. They utilized the buoyancy of the light molecules in the region of high temperature, and reversely, the tendency of the heavy atoms or molecules to settle down in cooler regions, in order to produce an increasingly effective separation. The Clusius separation tube consists of a vertical glass tube about 20 yd long and several centimeters in diameter. A wire, heated electrically to several hundred degrees centigrade, extends along the axis of the tube. The lighter components of the isotope mixture rise upward near the hot wire, whereas the heavier isotopes sink down near the cooler wall of the tube. The lighter isotopes can be drawn off from the top of the tube and the heavier isotopes from the bottom. To separate the isotopes completely, the gases drawn off can be processed again in another separation tube. With a permanent arrangement of several such tubes large quantities of isotopes can be obtained. Although this method works best with gaseous mixtures, it can also be used for separating liquid isotopes.

2-7. Photons

Up to now we have spoken only of the atomic constituents of matter and of their properties. However, one of the most important problems with which atomic physics deals is that of the interaction between matter and electromagnetic radiation. It is to be understood that by radiation we mean its manifestations in the whole wavelength region from the shortwave γ-radiation found in cosmic rays and nuclear physics through X-rays, ultraviolet, visible, infrared, and heat rays, up to and including what, in a restricted sense, are called electromagnetic waves. These various wavelength regions, which differ phenomenologically to a great extent, are plotted in Fig. 24 with the corresponding scales of wavelengths, frequencies, and energies. This last scale will be discussed below.

The transition from classical to atomic physics is intimately linked with the transition from continuous to discrete atomic phenomena. This is true for radiation as it is for matter. Classical physics is concerned always with light waves which are continuous in space and time; for example, a spherical wave emitted by a light source. As we shall see in the next chapter and throughout this book, the phenomena of atomic physics, such as the interaction between matter and radiation, can be explained only by assuming that radiation is emitted and absorbed as single quanta which we call photons. According to Planck's fundamental relation, the energy of these photons is proportional to the frequency $\nu = c/\lambda$ of the radiating dipole,

$$E = h\nu \tag{2-51}$$

h is a universal constant with the dimensions of action, energy × time. It is called Planck's constant and is equal to

$$h = (6.623 \pm 0.001) \times 10^{-27} \quad \text{erg-sec} \tag{2-52}$$

From (2-51) it follows that the energy of electromagnetic radiation is inversely proportional to its wavelength. This explains, for instance, why photographic plates react more to violet and blue light than to red light and why X- and γ-rays are more penetrating (and more dangerous to organisms) than longer-wave radiation which has less energy. Many other similar examples can be cited.

Next to the energy of photons we are most interested in their inertial mass and in their momenta, these being of importance for collisions between photons and atomic particles. Photons can have no rest mass because they do not "exist" unless traveling with the velocity of light. However, according to the mass-equivalence equation (2-36), every energy E has an equivalent mass E/c^2. Therefore, a photon of energy $h\nu$ has an inertia corresponding to a mass

$$m = h\nu/c^2 \tag{2-53}$$

and a momentum (mass × velocity c)

$$p = h\nu/c \tag{2-54}$$

The fundamental relation (2-51) was discovered by Planck in 1900, when he tried to derive theoretically his empirical relation for the spectral energy distribution of blackbody radiation. A blackbody radiator is one which completely absorbs incident radiation of every

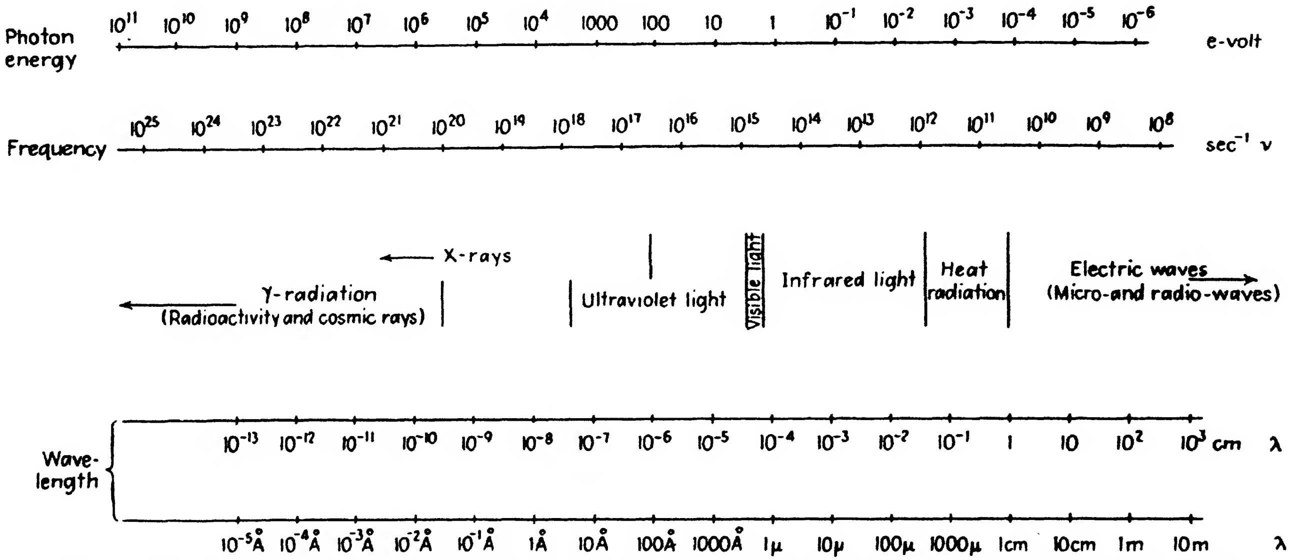


FIG. 24. Survey of the whole electromagnetic spectrum with wavelength, frequency, and energy scales (energy in electron volts).

wavelength. Such a radiator emits continuous radiation whose spectral energy distribution depends exclusively on the absolute temperature T, in a manner illustrated by Eq. (2-7) and by Fig. 25. Planck had originally obtained his formula, which agreed well with experimental results, by empirically modifying the less exact formula of Wien. The rigorous theoretical derivation of his formula (2-7), however, was not possible on the basis of the assumption, self-evident in classical physics, that the radiating dipoles ("resonators") can vibrate with any amplitude and energy. The correct relation (2-7) can be explained only on the basis of the revolutionary assumption that *the vibrating energy of a resonator is proportional to its frequency, Eq. (2-51), and always an integral multiple of $h\nu$.* This assumption of quantized energy states of a blackbody radiator is the origin of quantum physics. Some years later, Einstein applied Planck's formula to radiation itself and conceived the idea of the light quanta or photons. We know quite a number of phenomena which furnish evidence for the corpuscular nature of the photons. One of these, the photoelectric effect, has already been treated on page 31. Another pertinent experiment, the Compton effect, will be treated on page 194 and leads to a discussion of the complicated

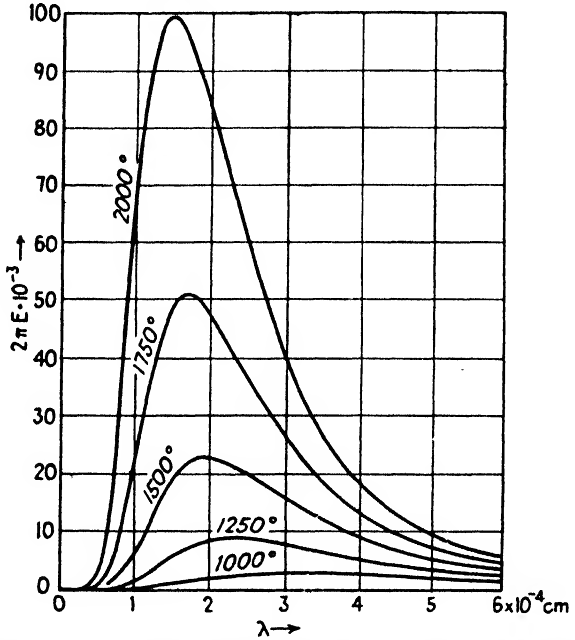


FIG. 25. The spectral energy distribution of a blackbody radiator at temperatures between 1000 and 2000° K, plotted against wavelengths.

question of the relation between the wave and quantum properties of radiation.

To sum up: In the realm of quantum physics, instead of a spherical wave being radiated from a source, we have an irregular statistical emission of photons into all directions of space. Only if we consider an average over a sufficiently long time do we arrive essentially at the classical result. Every surface element of a sphere surrounding the radiation source then receives an equal number of photons and thus an equal amount of energy. Impressive evidence of this emission of single photons is given by an experiment of Joffe. As a radiation source he used a very small anticathode of an X-ray tube, and as a receiver a small metal ball from whose surface photoelectrons were released by incident X-rays. The small collecting sphere was placed at such a distance from the source that its cross section was only about one-millionth of the surface of an imaginary sphere whose center was at the source. Now if the X-ray tube was excited to emit very short pulses of X-rays, it was found that it required about 10^6 pulses to photoelectrically release an electron from the small receiver sphere; the overwhelming majority of the emitted photons do not strike the sphere! This is a very striking demonstration of the corpuscular nature of photons.

Literature

General Literature

Briegleb, G.: Atome und Ionen, vol. 2/IA, Hand- und Jahrbuch der Chemischen Physik, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1940.

Harnwell, G. P., and J. J. Livingood: Experimental Atomic Physics, McGraw-Hill Book Company, Inc., New York, 1933.

Ruark, A. R., and H. C. Urey: Atoms, Molecules and Quanta, McGraw-Hill Book Company, Inc., New York, 1930.

Stranathan, J. D.: The Particles of Modern Physics, The Blakiston Company, Philadelphia, 1942.

SECTION 4:

Brueche, E., and A. Recknagel: Elektronengeraete, Springer-Verlag, Berlin, 1941.

Hughes, A. L., and L. A. DuBridge: Photoelectric Phenomena, McGraw-Hill Book Company, Inc., New York, 1932.

Klemperer, O.: Electron Optics, Cambridge University Press, London, 1939.

Massachusetts Institute of Technology, Radiation Laboratory Series, McGraw-Hill Book Company, Inc., New York: vol. 1, Radar System Engineering, 1947; vol. 2, Radar Aids to Navigation, 1947; vol. 3, Radar Beacons, 1947; vol. 4, Loran, 1948; vol. 5, Pulse Generators, 1948; vol. 6, Microwave Magnetrons, 1948; vol. 7, Klystrons and Microwave Triodes, 1948; vol. 8, Principles of Microwave Circuits, 1948; vol. 9, Microwave Transmissions Circuits, 1948; vol. 10, Waveguide Handbook (in press); vol. 11, Technique of Microwave Measurements, 1947; vol. 12, Microwave Antenna Theory and Design, 1949; vol. 13, Propagation of Short

Radio Waves (in press); vol. 14, Microwave Duplexers, 1948; vol. 15, Crystal Rectifiers, 1948; vol. 16, Microwave Mixers, 1948; vol. 17, Components Handbook, 1948; vol. 18, Vacuum Tube Amplifiers, 1948; vol. 19, Waveforms, 1948; vol. 20, Electronic Time Measurements, 1948; vol. 21, Electronic Instruments, 1948; vol. 22, Cathode Ray Tube Displays, 1948; vol. 23, Microwave Receivers, 1948; vol. 24, Threshold Signals (in press); vol. 25, Theory of Servomechanisms, 1947; vol. 26, Radar Scanners and Radomes, 1948; vol. 27, Computing Mechanisms and Linkages, 1948.

Millikan, R. A.: Electrons, Protons, Photons, Neutrons and Cosmic Rays, 3d ed., Chicago University Press, 1947.

Millmann, J., and S. Seely: Electronics, McGraw-Hill Book Company, Inc., New York, 1941.

Myers, L. M.: Electron Optics, Chapman and Hall, London, 1939.

Section 6:

Aston, F. W.: Mass Spectra and Isotopes, 2d ed., Longmans, Green & Co., Inc., New York, 1942.

Jones, R. C., and W. H. Furry: Separation of isotopes by thermal diffusion, *Rev. Mod. Phys.*, **18**, 151 (1946).

Mattauch, T.: Massenspektroskopie, *Ergeb. d. exakt. Naturwiss.*, **19**, 170 (1940).

Walcher, W.: Isotopentrennung, *Ergeb. d. exakt. Naturwiss.*, **18**, 155 (1939).

CHAPTER 3

ATOMIC SPECTRA AND ATOMIC STRUCTURE

"Ever since the discovery of spectral analysis, no one could doubt that the language of the atom could be understood if one learned how to interpret atomic spectra. The tremendous amount of spectroscopic data accumulated during the past 60 years seems at first to be too diverse and too complex to be disentangled. Seven years of X-ray spectroscopy have contributed even more to a clarification because here the problem of the atom is attacked at its root and the interior of the atom is revealed. What we hear nowadays of the language of the spectra is a true 'music of the spheres' of the atom, chords of integral proportions, an increasing order and harmony in spite of all diversity. For all times the theory of spectral lines will bear the name of Bohr. Yet another name will be permanently linked with his, the name of Planck. All integral laws of spectral lines and atomistics stem, originally, from the quantum theory. This is the mysterious organ on which nature plays her music of the spectra and according to whose rhythms the structure of the atom and the nucleus are arranged." Sommerfeld wrote these words in 1919 in the preface of the first edition of his famous work "Atombau und Spectrallinien." Since after 30 years they have lost nothing of their significance and beauty, they may be placed at the beginning of the chapter on atomic spectra and atomic structure. They characterize, at the same time, in a manner unequaled, the fundamental importance of spectroscopy for atomic physics. Therefore, we begin our discussion with a survey of the methods of spectroscopy.

3-1. Recording, Evaluation, and Classification of Spectra

a. Spectroscopic Apparatus and Their Application in the Various Spectral Regions

The spectroscopist is interested in the whole wavelength region in which spectra of atoms and molecules can be observed, i.e., the wavelength region between 10^{-9} and 100 cm. On page 68 and in Fig. 24 a survey of the different spectral regions is presented. Each of these spectral regions requires a special experimental technique of investigation. Wavelengths are usually measured in units of 10^{-8} cm, known as the angstrom unit and abbreviated A. However, in the infrared region the unit used is the micron $1\mu = 10^{-4}$ cm, and in the X-ray region the unit is the X-unit, $\text{XU} = 10^{-3}\text{ A} = 10^{-11}$ cm. Table 4 presents a brief survey of the various spectral regions and the spectroscopic equipment and detectors used in each region.

Table 4. The Different Spectral Regions and the Spectroscopic Equipment Used in Them

Wave length	Spectral region	Spectroscopic equipment	Radiation detector
$\lambda < 100$ A	X-ray	Crystal or grating spectrograph	Ionization chamber, photographic plate, Geiger counter
100–1800 A	Vacuum ultraviolet	Vacuum concave grating ($\lambda > 1250$A fluoride prism spectrograph)	Schumann photographic plate
1800–4000 A	Quartz ultraviolet	Quartz spectrograph or grating	Photographic plate
4000–7000 A	Visible region	Glass spectrograph or grating	Photographic plate
7000–10,000 A	Photographic infrared	Glass spectrograph or grating	Infrared-sensitive photographic plate
1–20μ	Shortwave infrared	Rock salt spectrograph or grating	Photoresistive cell, thermocouple bolometer, microradiometer, radiometer
20–400μ	Longwave infrared	Echelette grating	
$\lambda > 1000\mu$	Micro- and radio waves with special techniques		

The best known and most accessible spectral regions are the visible and ultraviolet between 7000 and 2000 A where simple spectrographs and fast standard photographic plates can be used. Glass optics are used in the visible region, 7000 to 4000 A. Because glass absorbs in the ultraviolet, quartz prisms and lenses have to be used in the 4000 to 2000 A region. A small, single-prism spectrograph (Fig. 26) is used to make a survey of the whole spectral region on one photographic plate. For detailed studies of a limited spectral region at high dispersion and high resolution, instruments with three or more prisms and a long focal length (Fig. 27) are used. For precision investigations, especially in the long wavelength region where the dispersion of optical media is

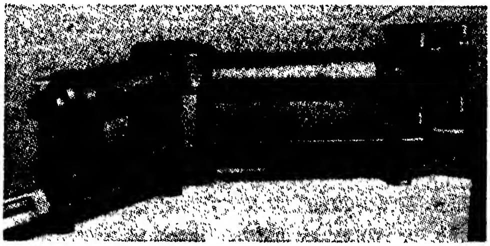


FIG. 26. A small single-prism quartz spectrograph manufactured by Zeiss.

small, large grating spectrographs are used. Figure 28 is a photographic picture of such a spectrograph using a Rowland concave grating.

In the spectral region below 1800 A the atmosphere begins to absorb and quartz ceases to transmit. Furthermore, the gelatin used in the

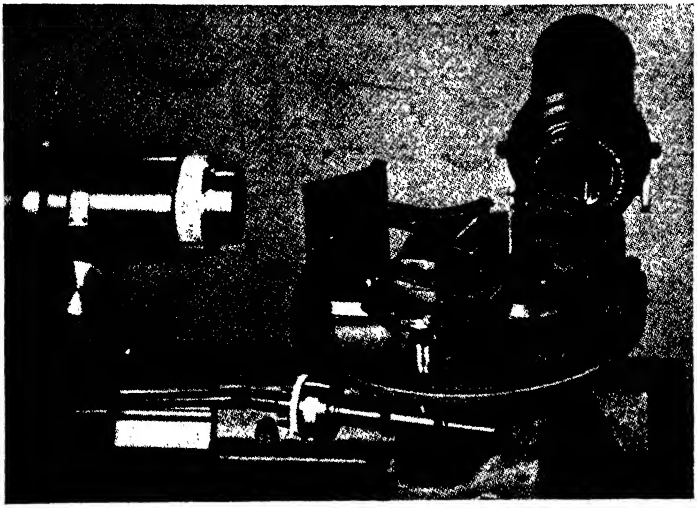


FIG. 27. The three-prism glass spectrograph of Zeiss (cover removed).

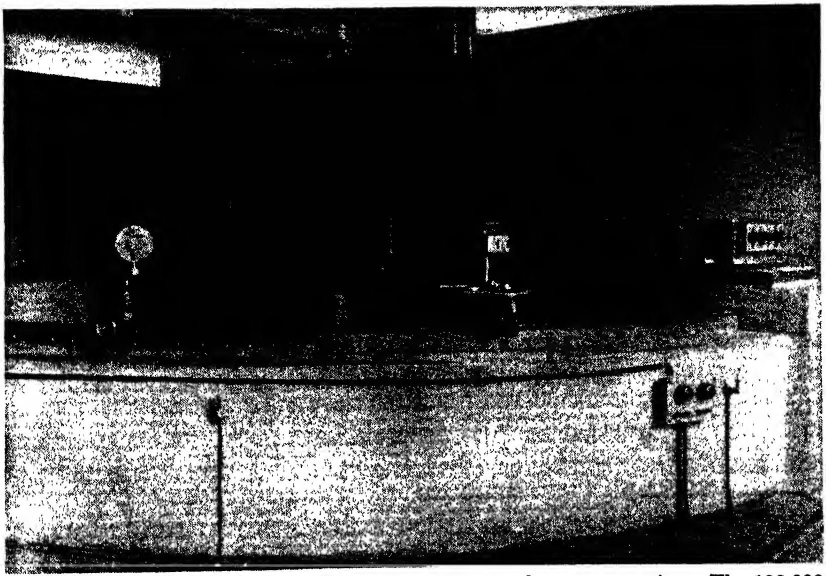


FIG. 28. Large grating spectrograph with 20 ft Rowland concave grating. The 120,000-line grating can be seen in the background on the left, two slits (for alternating work in different spectal regions and orders) in the foreground, plate holder in the right foreground. (Formerly mounted in the physical laboratory of the University of Bonn.)

ordinary photographic plate begins to absorb. For wavelengths below 1800 A, therefore, the whole apparatus including the light source and photographic plate must be placed in a highly evacuated chamber; consequently this region is called the vacuum ultraviolet. Besides, Schumann photographic plates, which do not contain gelatin in the

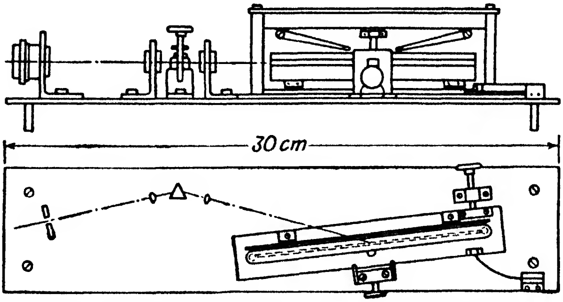


FIG. 29. Small fluorite vacuum spectrograph (without vacuum-tight cover) of Cario and Schmidt-Ott (schematic).

photographic layer, must be used to record the spectra. Since fluorite does not absorb down to 1250 A, a small vacuum prism spectrograph using fluorite optics can be used for recording an over-all survey of the spectrum in this region (Fig. 29). To obtain high dispersion in the short wavelength region down to X-rays, a concave grating ruled on speculum metal, which makes additional focusing optics unnecessary,

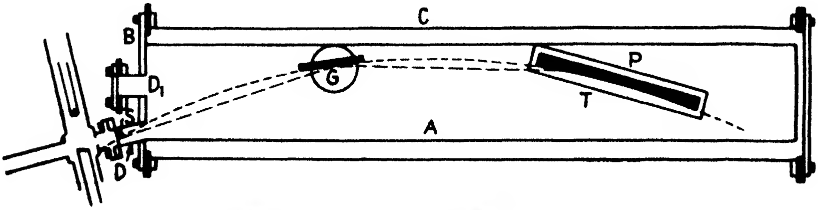


FIG. 30. Vacuum-grating spectrograph for spectroscopy in the extreme ultraviolet and longwave X-ray region. S = slit, G = grating, P = photographic plate. (*After Siegbahn.*)

is used. The grating as well as the source, slit, and plate must be used in a large vacuum chamber (Fig. 30). The grating constant of the optical gratings, mostly 10^{-4} cm = 1 μ, is too large for the shortwave vacuum ultraviolet. In this case the radiation is directed at the grating at a grazing incidence so that the effective grating constant is the projection of the actual grating constant on the direction of the incident light (Fig. 31). Because of the difficulty of adjusting vacuum instruments, much more experience is required to work in this region than in the visible or quartz ultraviolet regions. However, the results of this work have proved of utmost importance for the analysis of the structure of the atoms.

Interferometers such as the Fabry-Perot, Lummer-Gehrcke, etc., are required for high-dispersion and high-resolution work on the fine structure of spectral lines. However, these instruments are restricted to the visible and, by using quartz optics, the quartz ultraviolet regions.

The techniques for working in the photographic infrared region do not differ fundamentally from those in the visible, except that infrared-sensitized plates have to be used. The plates must be sensitized before exposure, and the exposure time increases the farther we want to reach out in the spectrum. The present limit for this technique is about 13,600 A, but extremely high exposure times are required to reach much beyond 11,000 A. Even here the exposure times are longer than those required in the visible region by two or three powers of ten. Glass spectrographs are used for a survey exposure in this region although the dispersion of glass at these wavelengths is very small. For detailed investigations in the infrared the grating is indispensable.

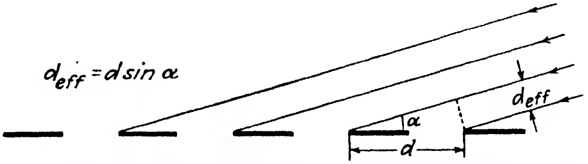


FIG. 31. Decrease of the effective grating constant at grazing incidence.

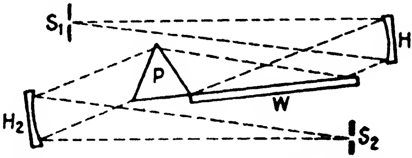


FIG. 32. Schematic arrangement of an infrared prism spectrograph with concave mirrors for collimating and focusing. A similar arrangement may be used with a grating instead of a prism. S_1 = collimator slit, S_2 = exit slit for monochromatic radiation, P = prism, W = plane mirror.

In the so-called shortwave infrared between 1 and 20 μ prism spectrographs are used with fluorite, rock salt, or other alkali-halide prisms having transmissions limits of 9, 16, and up to 40 μ, respectively. Silvered or gold-plated concave mirrors are used instead of lenses to focus the radiation, and thus absorption losses and chromatic aberration are avoided (Fig. 32). In the longwave infrared a grating, especially the so-called eschelette grating, is used. These have much larger grating constants than are used in the visible region. With suitably shaped grating lines almost all the diffracted radiation can be directed into *one* order. In this manner the superposition of various orders of the spectrum is avoided and, at the same time, more energy is made available in the one order. This is an important advantage since the energy available in the spectrum is rather small with a grating compared with a prism spectrograph.

Until recently photographs could not be made of the infrared spectrum beyond 1.4 μ. Czerny developed a method in which an infrared spectrum was projected upon a thin oil film which evaporated according to the energy in the different parts of the spectrum. The relief structure of the oil film is photographed by shining visible light on the film to form an interference pattern where the film has been evaporated. In the customary methods the intensity in the infrared spectrum is measured point by point. In the shortwave infrared, photoresistive cells and photoelements can be used as receivers, but in the rest of the infrared region thermopiles, thermocouples, and bolometers are used with either galvanometers or amplifier systems to record the spectrum.

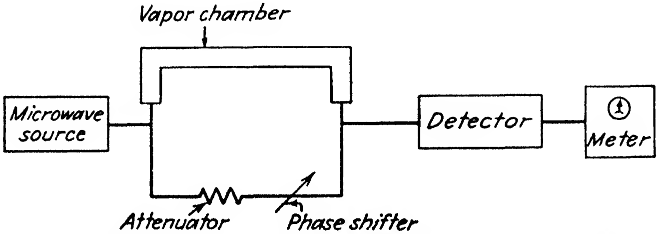


FIG. 33. Schematic diagram of microwave bridge spectrometer. (*Courtesy of M. W. P. Strandberg.*)

Only very recently has the spectral region of the longest wavelengths been opened for investigation by the new radio- and microwave spectroscopy. Only absorption methods can be used. In order to indicate one of the many possible methods for measuring microwave absorption, Fig. 33 presents a schematic diagram of a bridge spectrometer. The electric waves which are emitted by a calibrated high-frequency generator (microwave source in Fig. 33) are modified with respect to amplitude and phase in the absorption tube (vapor chamber). This change is compensated by means of the attenuator and phase shifter, until the meter on the right is at zero. The whole absorption curve can then be determined from the readings of the attenuator if the source frequency is varied by small steps. Practically, a frequency modulation method is used in most cases; with it the whole absorption curve can be made visible at once on the screen of an oscilloscope. A dispersing element (prism or grating) is not required in a microwave spectroscope, because the source itself provides monochromatic radiation of variable frequency. The sensitivity of this method, by which even the absorption of only 10^{16} atoms can be measured, can apparently be increased considerably. The accuracy with which a frequency can be measured in this region of extremely long wavelengths is unsurpassed. Since a klystron can radiate only in a narrow wave band, only one line or a group of very closely spaced lines can be studied at one time. The regions near 1 and 3 cm have been investigated most extensively.

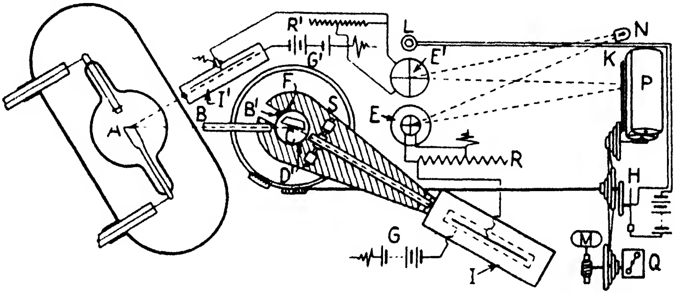


Fig. 34. X-ray spectrograph with rotating crystal C, ionization chamber I, and recorder for the dispersed radiation of the anticathode A. In order to control the constancy of the radiation, the undispersed radiation is recorded simultaneously. (*After A. H. Compton.*)

A special experimental technique is required also for the extreme shortwave region, that of the X-rays (0.01 to 100 A). There is no substance which can be used for prisms or lenses to refract X-rays. However, ruled concave gratings with grazing incidence can be used, just as in the far ultraviolet (Fig. 30). Figure 67 shows a beautiful example of such an X-ray grating spectrum. Moreover, since von Laue discovered, in 1912, that X-rays are diffracted by the atoms in crystal lattices, crystals with regularly spaced atoms (see page 425) are used in crystal grating spectrometers. A crystal, however, is not a *line* grating, but a three-dimensional *point* grating. For a given crystal orientation relative to the direction of incident radiation, therefore, there corresponds to each wavelength a definite angle, of which this wavelength can emerge from the crystal. In order to photograph the *whole* spectrum, according to Bragg, the crystal and the receiver must be slowly rotated in such a way that the bisector of the angle between incident and recorded beam is always perpendicular to the crystal plane (Fig. 34). Figure 35 shows an X-ray spectrum recorded this way. As the microcrystals in a crystal powder have all possible orientations, monochromatic X-rays, according to Fig. 36, will be diffracted by a powder sample, *Pr*, through angles which depend on the wavelengths of the X-rays and the lattice constants of the microcrystals. This method of Debye and Scherrer thus permits measuring (from a ring pattern such as Fig. 37) the wavelength of the X-rays if the lattice constants are known or (more commonly applied this way) to determine the lattice constants of the microcrystals if the wavelength is known.

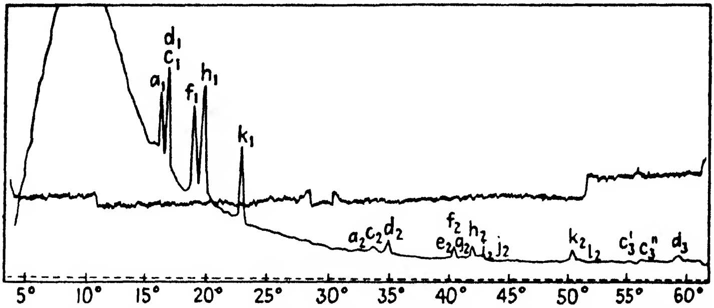


FIG. 35. The X-ray spectrum of a tungsten anticathode (continuous retardation spectrum with superposed X-ray lines of *W*), recorded with the instrument shown in Fig. 34.

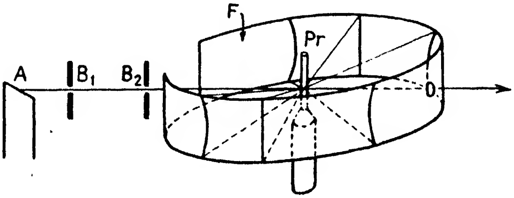


FIG. 36. Schematic Debye-Scherrer arrangement for photographic recording of the X-ray diffraction by crystal powders. *A* = anticathode of the X-ray tube, B_1 and B_2 = diaphragms, *Pr* = sample of crystal powder, *F* = photographic film.

FIG. 37. X-ray diffraction pattern of CeO_2 crystal powder taken with method illustrated in Fig. 36.

b. Emission and Absorption Spectra

We distinguish between emission and absorption spectra according to whether the atom under consideration itself emits the wavelength in question, or absorbs it out of the continuous wavelength range of the incident radiation (continuous radiation). In the first case, under direct observation, the spectrum appears bright on a dark background, in the second case it is dark on a bright background. Examples of emission line spectra are the spectra of all incandescent gases (Fig. 38), while the best known absorption line spectrum is that of the Fraunhofer lines in the sun's spectrum. These lines, which are a result of the absorption of gas atoms in the sun's atmosphere, appear dark against the bright continuous emission spectrum of the lower layers of the sun, the photosphere. Figure 41 is an example of an atomic absorption spectrum. A special case of an absorption spectrum is the so-called

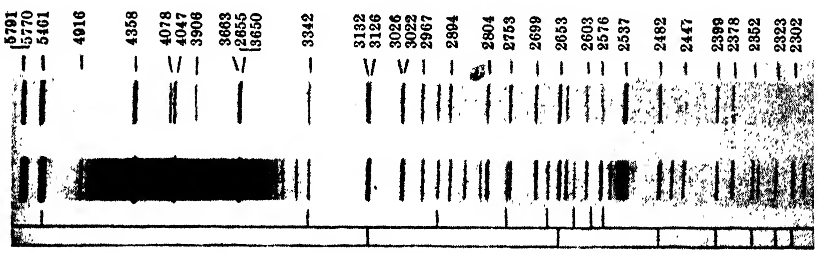


FIG. 38. Atomic line spectra of mercury: upper spectrum normally exposed, lower spectrum overexposed so that weaker lines and the visible continuous spectrum of the Hg_2 molecule become apparent (see page 385 and Fig. 164*d*). Below the spectrum, two line series are indicated which are superimposed in the spectrum. (*After Herzberg.*)

self-absorption of spectral lines. This phenomenon is based on the fact that an atom, as we shall see on page 102, can absorb many of the lines which it emits. Since, according to page 187, the width of a spectral line depends on the disturbance of its emitting or absorbing atoms by its environment, the Cu lines emitted by a copper spark are broader than the Cu lines absorbed by the cooler copper vapor of the outer envelope of the spark. The Cu atoms of the outer envelope thus absorb, so to speak, the center of the broader Cu emission line, so that a sharp absorption line is seen against the background of the broader emission line. Figure 39 is an example of a density curve of a pure emission line of copper and the same line "self-reversed" because of self-absorption.

To observe an emission spectrum, the atoms under investigation must be "excited." We shall return to the physical meaning of an excited atom on page 87. In the X-ray region it results from fast electrons (cathode rays) impinging on the anticathode; in the optical region it is the result of the electron impact in a gas discharge or of thermal collisions at sufficiently high gas or vapor temperatures. Glow discharges, arc and spark discharges of various types, as well as flames are used to excite optical spectra. Arcs and sparks, because of their high temperatures, also serve to vaporize solid materials, for instance of the electrodes, and then to excite the atoms of the vapor. While we thus produce excitation mostly by electrical or chemical methods, we find that there is purely thermal excitation in the atmospheres of the sun and fixed stars.

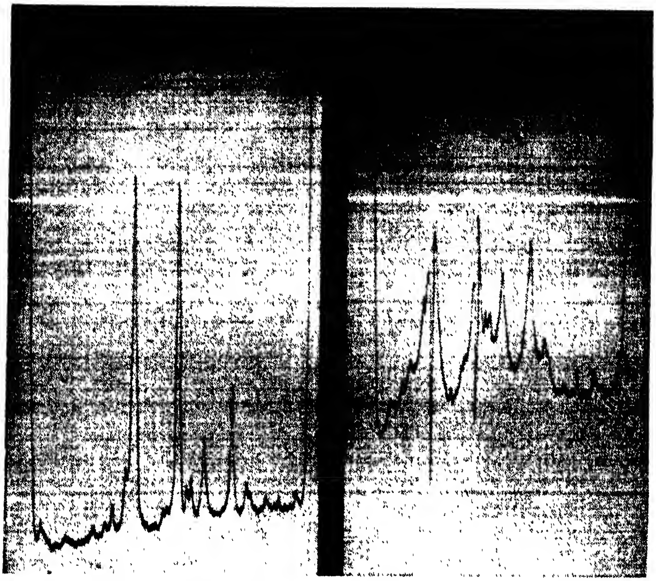


FIG. 39. Densitometer record as an example for the evaluation of a spectrum by means of a recording microphotometer. Of the four most intense Cu lines, the two on the left are resonance lines which appear self-absorbed on the right side.

To record an absorption spectrum (Fig. 41) it is necessary to have a continuous emission spectrum, i.e., one which contains all wavelengths in the spectral region under consideration. The absorbing substance (e.g. gas atoms), which absorbs its characteristic spectral lines, is placed between the source, the continuous spectrum, and the spectrograph (Fig. 40). In the visible and infrared regions an incandescent solid such

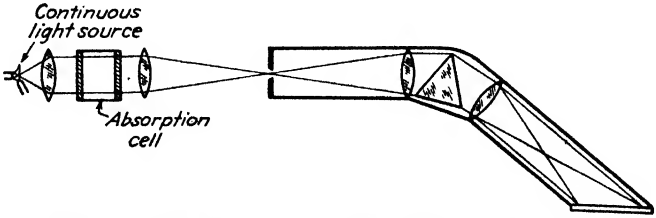


FIG. 40. Arrangement for photographing absorption spectra.

as a tungsten filament lamp or the positive crater of a carbon arc are used as continuous emission sources. In the ultraviolet, the continuous spectra of various discharges, especially the hydrogen molecular continuum, are used; in the shortwave ultraviolet, the helium molecular continuum (page 387). We shall deal with the X-ray absorption spectra later (page 120) because this is a more complicated problem.

FIG. 41. Spectral line series of the sodium atom absorbed by sodium vapor. (*After H. Kuhn.*)

c. Wavelength and Intensity Measurements

The first step in evaluating a line spectrum consists of determining the wavelengths of the lines as well as (at least roughly) their relative intensities. On the other hand, the characteristic feature of a continuous spectrum is that its intensity is a function of the wavelength, and it is the intensity distribution which must be determined. During the last 70 years, by far the greatest part of all spectral lines, at least of the atomic lines, has been measured, so that for most purposes it is sufficient to identify spectral lines rather than measure their wavelengths. For this purpose a comparison spectrum is photographed on the same plate with the spectrum being investigated. The comparison spectrum is usually the spectrum of mercury (Fig. 38) or of iron (Fig. 76), both of which have been carefully and accurately measured. The wavelengths of the lines of the unknown spectrum are determined either by interpolating between the lines of the comparison spectrum or by drawing a dispersion curve, using the comparison spectrum for calibration. To get the dispersion curve of a spectrograph, the distances of the lines from one reference line are plotted against the corresponding wavelengths. The distances of unknown lines are measured under a microscope, or, less accurately, under a spectrum projector, and the corresponding wavelengths then taken from the dispersion curve.

As we have mentioned, wavelengths are usually given in angstrom units, $A = 10^{-8}$ cm, and in the infrared in microns, μ. Because of the relation $\lambda = c/\nu$, where c is the velocity of light in the medium under consideration, the wavelengths depend on the refractive index of the medium in which the light is propagated. In order to have exactly comparable values, the measured wavelengths must be reduced to those they would have in vacuum. As we shall soon see, the laws of line spectra are simplest not in terms of wavelengths, but in terms of frequencies, $\nu = c/\lambda$ or wave numbers, $\bar{\nu} = 1/\lambda$ (the number of waves per centimeter). This computation of $1/\lambda$ is usually combined with the reduction to the wavelength in vacuum. H. Kayser has carried out these computations and published them in his "Tabelle der Schwingungszahlen," from which for each wavelength measured in air the corresponding vacuum-reduced wave number $\bar{\nu}$ cm^{-1} can be taken.

Line intensities are mostly determined by comparison according to an arbitrary scale, in which the darkest (strongest) line is graded as 10 and the weakest, barely recognizable line as 0. For more accurate investigations, the density of the photographic plate at the location of a line is measured photoelectrically or thermoelectrically with a densitometer. With a recording instrument the distances between lines, line intensities, and the shape of each line (line contours) of a spectrum can be measured simultaneously. A typical curve obtained from these measurements is shown in Fig. 39. In making absolute measurements and in comparing the intensities at various wavelengths, it must be noted that the density of the plate and the intensity of a line are not proportional. The relation between the two must be determined from the gradation curve of the particular emulsion used. For absolute intensity measurements, especially of continuous spectra, a comparison spectrum of known intensity distribution (calibrated tungsten lamp or carbon arc crater) is compared with the spectrum being investigated. This comparison, and thus the measurement of the intensity distribution, can be made by photographic photometry. The prevalent method, however, consists of dispersing the spectrum in a double monochromator and measuring or recording the intensities by amplifying the output of a photoelectric cell, thermocouple, or bolometer, and, occasionally, also by measuring this output directly with a galvanometer or electrometer.

d. Lines, Bands, and Continuous Spectra

Very early, spectra had been classified according to their appearance as line, band, or continuous spectra. Figure 42 shows examples of the three types. Much later it was realized that the distinction between the first two groups is directly related to their origin; line spectra are always emitted or absorbed by atoms or atomic ions, while the emitters of band spectra are always molecules. We shall deal in some detail with band spectra in Chap. 6. Line spectra were further classified as arc or spark spectra according to the source in which they occurred with highest intensity. A more precise terminology was introduced when it became known that arc spectra belong essentially to neutral atoms, whereas spark spectra were emitted by ionized atoms, or ions. The term arc spectrum is now understood to mean the spectrum of a neutral atom and is labeled by I, so that Fe I designates the arc spectrum of iron. In the case of spark spectra the terms first, second, etc., spark spectrum of iron are used. The spectra are classified according to the ions, Fe^+, Fe^{++}, etc., which emit them and are labeled Fe II, Fe III, etc. Thus the Fe IV spectrum is the third spark spectrum of iron and is emitted by the Fe^{+++} ion.

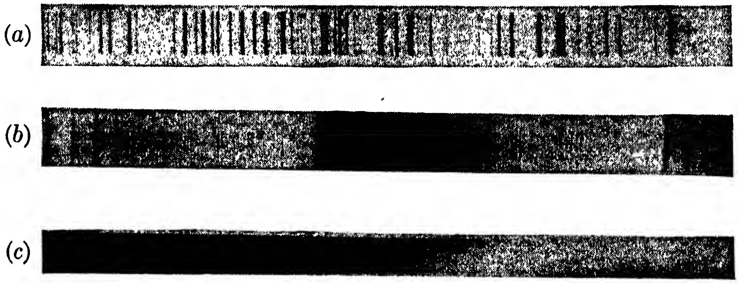


FIG. 42. Examples of (*a*) an atomic line spectrum, (*b*) a molecular band spectrum, and (*c*) a continuous spectrum.

With respect to the third group, the continuous spectra, a plain classification as to their origin is not possible without a detailed investigation. As we shall see, atoms as well as molecules can emit continuous spectra but they differ clearly in their excitation conditions. Into this class of the atomic spectra belongs also the important continuum which is emitted when free electrons collide with ions (electron retardation continuum or, using the German term, bremscontinuum). We shall return to it on page 110. Continuous spectra are also emitted by nearly all incandescent solid bodies and they all show a clear relation to the continuous blackbody spectrum which we discussed on page 64. The total radiation, the spectral energy distribution of the radiation, and the wavelength of the intensity maximum of the blackbody continuum can be determined from Planck's law which also contains the older laws of Stefan-Boltzmann and Wien.

3-2. Series Formulas and Representation of Terms of Line Spectra

An inspection of a simple line spectrum such as that of the hydrogen atom (Fig. 43) reveals that there must exist a sequential relationship between the wavelengths of its spectral lines. A regular sequence of lines in a spectrum is called a series, and the order in which the wavelengths of the lines of a series occur in a sequence is called a series rule.

The first series rule was formed in 1885 by Balmer, a high-school teacher in Basle, for the first four lines H_α, H_β, H_γ, and H_δ of the hydrogen series. This series lies in the visible and near ultraviolet and is named after him. In the following years Kayser and Runge analyzed the accumulated mass of wavelength data of all available atoms which had been measured with great precision by them and their coworkers. They deduced a great number of series rules and showed that, in addition to hydrogen, all atomic spectra consist of series of lines in the visible and ultraviolet regions, the only spectral regions which they were able to investigate. Furthermore, Kayser and Runge realized that it was not appropriate to use the wavelength as a function of some parameter, as Balmer had done; they used the wave number, $\bar{\nu} = 1/\lambda$, that is the number of waves in one centimeter of light path in a vacuum, measured in reciprocal centimeters, cm^{-1}. Rydberg finally found the correct form for the series rule concealed in Balmer's formula. In this form the wave numbers of the spectral lines are written as the difference of two quantities, which he called "terms." The physical significance of these terms, whose dimensions are the same as that of wave numbers, reciprocal centimeters, will become clear later. With the Rydberg modifications the Balmer formula for the visible hydrogen series becomes

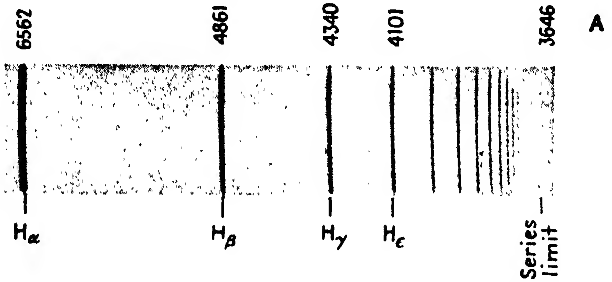


FIG. 43. The Balmer-line spectrum emitted by the hydrogen atom. (*After Herzberg.*)

$$\bar{\nu} = \frac{R}{2^2} - \frac{R}{n^2} \qquad n = 3, 4, 5, \ldots \tag{3-1}$$

Here R is a constant with the dimension cm^{-1}, the so-called Rydberg constant, and n is the series number of the second term. We shall see in the next section that the representation of a spectral line as a difference between two terms has considerable theoretical significance and we shall learn at the same time the physical meaning of a "term." The accuracy with which the simple Balmer formula (3-1) can describe the wave numbers and thus the wavelengths of the hydrogen spectrum is shown in Table 5, where measured and computed values of the first 13 lines of the Balmer series (this includes lines not known to Balmer) are compared. This table also illustrates the accuracy with which spectroscopic measurements can be carried out, though in this case, because of line width, the values are relatively inaccurate.

Table 5. Comparison of Computed and Observed Wavelengths of Several Lines of the Balmer Spectrum of the H *Atom*

Name of line	n	Observed, A	Computed, A
H_α	3	6563.07	6563.07
H_β	4	4860.90	4861.57
H_γ	5	4340.47	4340.63
H_δ	6	4101.73	4101.90
H_ϵ	7	3970.07	3970.22
H_ζ	8	3888.99	3889.20
H_η	9	3835.35	3835.53
H_ϑ	10	3797.91	2798.04
H_ι	11	3770.78	3770.77
H_κ	12	3750.25	3750.36
H_λ	13	3734.54	3734.51
H_μ	14	3722.04	3722.08
H_ν	15	3712.14	3712.11

According to Rydberg, a series of spectral lines is represented by the general formula

$$\bar{\nu} = \frac{R}{(m+a)^2} - \frac{R}{(n+b)^2} \qquad n > m \tag{3-2}$$

a and *b* are characteristic constants of a particular series, *m* is a varying small integral number, characteristic for the various series of the spectrum, and *n* again the series number. The first term is called the fixed term and the second, the variable term of the series. The physical significance of all these constants becomes clear only on the basis of the quantum theory; we shall consider them in the next section.

An important step forward was accomplished when Ritz in 1908 formulated his combination principle. It states that by additive or subtractive combination of frequencies of spectral lines (or their corresponding terms) new spectral lines or terms of the atom under consideration are obtained. The Ritz combination principle led to the discovery of a large number of new spectral lines and so contributed very much to our knowledge of spectroscopy, although actually not *every* combination of two observed lines can be observed. All these relations which at first appeared to be rather mysterious will be explained in the next section on the basis of the Bohr theory. Only then shall we understand their full significance.

We know already that the H atom as a source of the Balmer spectrum according to (3-1) is characterized by a series of term values

$$T_n = \frac{R}{n^2} \quad cm^{-1} \qquad n = 1, 2, 3, \ldots \tag{3-3}$$

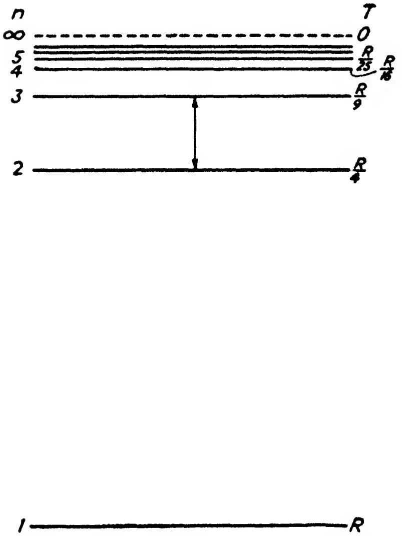


FIG. 44. Term diagram of the hydrogen atom with principal quantum numbers n (on the left), term values (on the right), and an arrow to indicate a transition between two terms.

from which by subtractive combination according to the combination principle the wave numbers of the H-atom lines (3-1) are obtained. For the following theoretical treatment of the spectra on the basis of the atomic theory a diagram representing the terms of an atom has proved very valuable. We shall call it the term diagram of the atom. Each term is represented in the diagram by a horizontal line and these lines are arranged in the order of increasing term number n, as shown in Fig. 44. Since the term value T_n is inversely proportional to the term number n, the term values increase in going from the top to the bottom of the diagram. For this reason they are called *negative* term values. The term value T_∞ corresponds to $n = \infty$, and the largest term value $T_1 = R$ corresponds to $n = 1$. Figure 44 was constructed in this manner from the terms (3-3) of the H atom. The series of term values converges with increasing term numbers n toward a limit. The physical significance of this limit will be explained in the next section. We mentioned already that by the Rydberg formula the wave number of each spectral line is expressed as the difference of two terms of the atom. This is illustrated in Fig. 44 by the arrow connecting the two terms under consideration, as in our case:

$$\bar{\nu}_{23} = T_3 - T_2 \tag{3-4}$$

It is the task of the spectroscopist to "isolate" the term diagram of atoms from the measured wavelengths of their spectra after having calculated the wave numbers $\bar{\nu}$ from them; this is called empirical term analysis in contrast to the theoretical term interpretation which we shall consider later. We shall discuss the isolation of terms with the aid of Fig. 45 in which a series of terms has been arbitrarily numbered 1, 2, 3, . . . , 8. Between these terms a number of transitions have been indicated. They are represented by the arrows *a*, *b*, *c*, . . . , *k*, *l*. Each arrow corresponds to a definite measured wavelength and its calculated wave number, so that the length of an arrow is a measure of the wave number $\bar{\nu} = 1/\lambda$.

From Fig. 45 it can be seen that by computing the wave number difference between two pairs of lines having the same upper state, for example *b* and *e*, the difference between two lower terms, in this case 1 and 2, can be determined. The wave-number difference between the lines *c* and *f* or between *d* and *g* gives the same term difference 1 − 2. By thus computing wave-number differences of lines, the distance between the two lower terms can be determined. Likewise, by computing the wave-number differences $e - h = f - i = g - k$ the term difference 2 − 3 can be obtained. Similarly, by computing the differences of wave numbers for lines from the same lower state, the higher term can be determined, for example, $c - b = f - e = i - h =$ term difference 5 − 6, etc. The empirical term analysis of a spectrum thus begins with looking for constant wave-number differences between pairs of lines. The necessary agreement among the differences can be computed from the accuracy of the wavelength measurements. The survey of a complicated spectrum and the discovery of constant differences can be made easier by plotting the spectrum along a wave-number scale and measuring wave-number differences with a pair of compasses. Though term differences can be found in this manner, it is generally difficult to decide at first to which place in the term diagram a particular term difference belongs, that is, where it lies with respect to the ground state. The construction of the complete term diagram thus requires either a theoretical discussion or additional experimental evidence. Of great help are especially electron impact experiments from which the excitation energy, i.e., the height of the upper term of a line, can be determined exactly. We shall not go into the finer details of the empirical term analysis since they are of interest only to the spectroscopist.

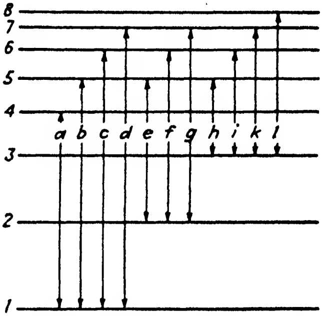


FIG. 45. Arbitrary example of a term diagram with transitions. Used in the text to explain the method of "isolating" terms.

3-3. The Fundamental Concepts and Postulates of the Bohr Theory

Since the work of Kayser, Runge, and Rydberg, it had become clear that, in a certain sense, the spectrum is a picture of the atom. There was no doubt that the spectrum is intimately associated with some kind of motion of atomic electrons rotating round a positive nucleus, and that the dynamic atomic model of Rutherford (page 23) was essentially correct. However, it was not until 1913 that the Dane Niels Bohr found the key which opened the way to an understanding of atomic spectra.

Bohr started from the Rutherford model in which the Coulomb attraction between the nucleus and the electrons was compensated by the centrifugal force due to the rotating electrons. Former theoretical attempts to explain the spectra had always been based on the classical concept that the nucleus and its electrons form electric dipoles and consequently must radiate. A relation between the frequency of rotation of an electron and the frequency of a spectral line, however, was sought for in vain. There were two main difficulties which no classical theory was able to solve: First, the existence of a stable atom could not be explained because the atom must continually radiate and thus lose energy as long as its electrons rotate. Second, there was no way to understand the emission of a spectrum of discrete frequencies with no apparent relation to the frequency of electronic rotation.

Bohr avoided these difficulties in a manner similar to severing the Gordian knot, by formulating his famous postulates which limited the validity of classical physics so far as the realm of atoms was concerned. These postulates, conceived with as much boldness as with admirable physical intuition, were not theoretically justified until 12 years later when Heisenberg and Schrödinger developed their quantum mechanics. We shall discuss this in detail in Chap. 4. Meanwhile Bohr had already proved in his first paper that his postulates were extremely useful for explaining atomic spectra. In the following years, the Bohr theory achieved success after success and won fame by explaining the periodic table.

What are these famous postulates? Bohr contended:

1. The electron can rotate around the positive nucleus only in orbits defined by certain quantum conditions. In these "quantized orbits," however, the electron rotates *without radiating,* contrary to the requirements of classical electrodynamics.

2. Each of these quantum orbits corresponds to a definite energy state E of the atom. The normal atom has its electrons rotating in

orbits of the smallest possible radius, i.e., with the lowest possible potential energy. In order to bring an electron into an outer orbit, a certain amount of energy, the "excitation energy" for the particular orbit, must be expended. After an average time of 10^{-8} sec in an orbit of higher energy, the electron "jumps" spontaneously to an orbit of lower energy and eventually to the ground orbit. The energy difference between the higher energy of the initial orbit E_i and the lower energy of the final orbit E_f is emitted as a spectral line of frequency ν given by Bohr's frequency condition

$$E_i - E_f = h\nu \tag{3-5}$$

where h again is the famous Planck constant (2-52).

According to Bohr's postulates only certain distinct orbits are possible for an electron. These "allowed" orbits are distinguished by a quantum condition: that the product of the circumference of the orbit $2\pi r$ and the momentum of the electron mv (a product which has the dimension of action) is equal to an integral multiple of Planck's constant h,

$$2\pi rmv = nh \qquad n = 1, 2, 3, \ldots \tag{3-6}$$

Thus Planck's constant enters into *both* of the fundamental equations of the Bohr theory (3-5) and (3-6), and therefore is characteristic for all atomic processes. Details will be presented in the theoretical treatment of the H atom on page 97.

By applying the Bohr postulates, it actually proved possible to interpret the atomic spectra without too much difficulty. By substituting in the Bohr frequency condition (3-5) the wave number $\bar{\nu} = \nu/c$ for the frequency, and solving the equation for $\bar{\nu}$, we obtain

$$\bar{\nu} = \frac{E_i}{hc} - \frac{E_f}{hc} \tag{3-7}$$

By a comparison with the Rydberg series formula (3-2)

$$\bar{\nu} = T_2 - T_1 \tag{3-8}$$

the significance of the "terms" T in the series formula becomes evident. *A term is equal to an energy state of a particular electron orbit divided by hc.* The term diagram shown in Fig. 44 thus can be regarded as an energy-level diagram. Each term value can be converted to an energy value by multiplying it by hc. Such an energy-level diagram represents the possible energy states of the quantum orbits of an atom. We shall show when treating the H atom that the quantum number n associated with the quantum condition (3-6) is identical with the term number n of the terms in Eq. (3-3).

We can characterize a certain atomic state with a quantum number n either by specifying its term value in reciprocal centimeters, with the numbers progressing from the highest term T_∞ to the lowest (Fig. 46), or by specifying its energy value $E_n = hcT_n$ in ergs. Instead of measuring the energy in ergs, it is customary in atomic physics to measure it in electron volts, the potential through which a charge $-e$ must fall, to acquire a kinetic energy E. This new energy unit is called an electron volt or ev and is defined by the equation

$$E = eV \tag{3-9}$$

These three energy units are related as follows:

$$1 \text{ ev} = 1.60 \times 10^{-12} \text{ erg}$$
$$= 8067.5 \text{ cm}^{-1} \tag{3-10}$$

We shall continually make use of these relations, and therefore they should be kept in mind. There is still another energy unit which is frequently used. Instead of dealing with the energy of a single atom it is possible to refer to the energy of 1 mole = 6×10^{23} atoms (page 18) and measure it in thermal energy units, kilocalories (kcal). This way of measuring energy in kilocalories per mole is generally preferred by chemists. The relation between this new energy unit and the electron volt (3-10) is

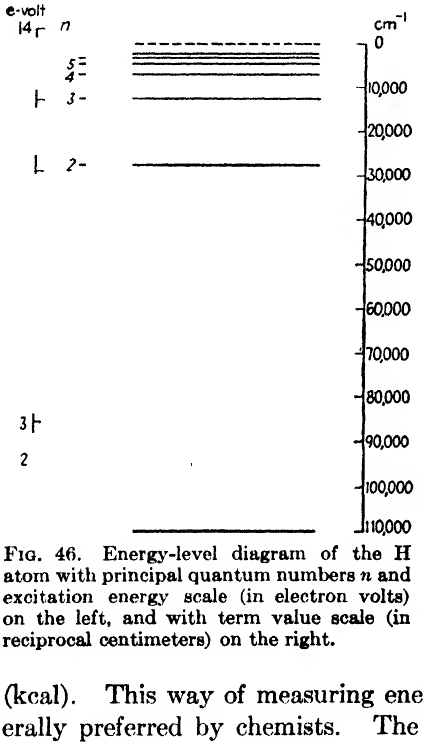


FIG. 46. Energy-level diagram of the H atom with principal quantum numbers n and excitation energy scale (in electron volts) on the left, and with term value scale (in reciprocal centimeters) on the right.

$$1 \text{ ev} = 23.04 \text{ kcal/mole} \tag{3-11}$$

After this digression into the different energy units which are important to atomic and molecular physics we return to the consideration of the energy level diagram, Fig. 46, and inquire into its physical significance.

For the normal atom the energy of the ground state is $-E_0$; the energy (as a binding energy of the electron) is taken to be negative. If an electron (we speak here of an H atom with only one electron) is now brought into an outer orbit, work must be done in order to overcome the electrostatic attractive force between electron and nucleus. By this transfer of external excitation energy the binding energy $-E_0$ of the atom is reduced. However, an atom cannot absorb any arbitrary amount of energy, but only those energies which correspond to differences between its own and any higher energy state in the energy level diagram. If we continue to transfer excitation energy to the electron and thus remove it to orbits farther and farther away from the nucleus, the consecutive excitation energies decrease more and more. The energy levels of the atom converge toward a limit where, in our notation, the energy of the electron is zero. This limit will be reached when the electron is no longer bound to its nucleus, or in other words, when it has been removed to infinity. *The term or energy value zero thus corresponds to the state of ionization where the electron is not only excited, but completely removed from the nucleus. From this consideration we see that the negative term or energy values correspond to the binding energy between the electron and the nucleus.* In the ground state of the normal atom, where the electron occupies the innermost orbit, the electron is bound most strongly to the atom, i.e., the atom has its greatest negative energy value. This relationship should be understood very clearly in order to avoid any confusion over negative terms and energy values. Actually, the relationship can be reversed and this is frequently done. If we start with the normal atom as it is most frequently found in nature, and denote the energy of its ground state by zero, we can let the energy values of its energy levels be positive (left scale in Fig. 46). In this case we speak of excitation energies. In other words, the negative energies are binding energies, and the positive energies are excitation energies. State E_2 in Fig. 46 can be regarded as having an excitation energy of 10.15 ev, or a binding energy of -3.39 ev, corresponding to a term value $-27{,}419\ \text{cm}^{-1}$. Computations involving excitation energies have considerable physical significance. *Only after we have transferred to an atom a certain excitation energy can a transition from an excited level to the ground level occur, in in which a spectral line of frequency*

$$\nu = \frac{E_i - E_f}{h} \tag{3-12}$$

is emitted.

Since, according to the Ritz combination principle, each difference between two energy levels in the energy level diagram corresponds to a spectral line (we shall consider exceptions later), it can be seen from Fig. 46 that an atom in its first excited level can emit only *one* line and that by increasing the excitation energy gradually the entire spectrum can be emitted. The Bohr theory thus indicates that there is a direct relationship between the excitation energy and the spectrum emitted by each atom. We shall consider the proof of this in the next section.

We also know that by exciting the electron of an atom to higher and higher energy levels, the electron is eventually separated from the atom. In other words, the excitation energy of the term limit E_∞ is the ionization energy of the atom. The proof of this contention of the Bohr theory we shall also present in the next section. Since the excitation and ionization energies are usually given in electron volts, the energy states computed according to (3-7) and given in volts are called "critical potentials."

In concluding this section, we return to a brief survey of Bohr's concepts and assumptions. Bohr declared classical electrodynamics invalid in so far as it required the electron to radiate when rotating in a "stationary" orbit. He correlated the change in energy of the atom, which occurs when an electron "jumps" from one orbit to another, with the emission or absorption of radiation, by means of his frequency condition (3-5). This was a departure from the old idea that the frequency of the emitted or absorbed radiation must be related to the rotational frequency of the electron. The Bohr theory does not, in general, account for any direct relationship between the rotational frequency of the electron around the nucleus and the emitted or absorbed frequency. However, we shall show on page 165 that for the limiting case of large quantum numbers and small energy changes the classical and quantum-theoretical predictions become identical. The Bohr theory makes no assertions about the mechanism of the emission or absorption of radiation which occurs together with electron jumps. It only states the relation between changes in energy states of the atom and the wavelength of the emitted or absorbed spectral line.

Before we apply the Bohr theory to the hydrogen atom, we shall present the experimental confirmation of the theory by the experiments of J. Franck and G. Hertz.

3-4. Quantum Jumps Due to Collisions

Bohr's frequency condition

$$E_i - E_f = h\nu = hc\bar{\nu} = \frac{hc}{\lambda} \tag{3-13}$$

implies the existence of a relation between the excitation energy E_i of an atom and the emission of a spectral line of wavelength λ. After Rau's preliminary experiments, J. Franck and G. Hertz in 1914 obtained conclusive experimental proof of the above consequence of the Bohr theory. A schematic diagram of their apparatus is shown in Fig. 47. Electrons emitted by the thermionic cathode *K* are accelerated by a variable potential *U* between *K* and the grid *G* of a vacuum tube filled with mercury vapor. In the space between *K* and *G* these electrons collide with gas atoms, e.g., Hg atoms. Electrons arriving at *G* pass through the grid and approach the collector electrode *A* which is charged negatively to about 0.5 volt. If the energy of the electrons is greater than 0.5 volt they arrive at *A* and pass through the galvanometer circuit to ground. Those electrons which have lost all or a large part of their energy in a collision with an Hg atom cannot overcome the opposing potential of 0.5 volt. Consequently, they do not arrive at *A* and are not measured by the galvanometer.

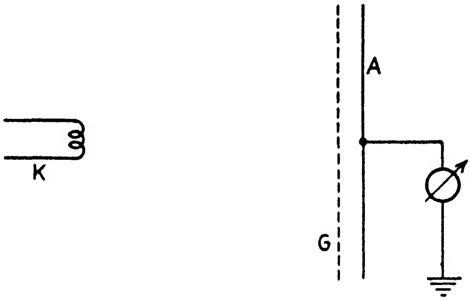


Fig. 47. Simplest arrangement of Franck and Hertz for detecting discrete stationary energy states of atoms by controlled electron collisions. *K* = thermionic cathode, *G* = accelerating grid, *A* = receiving electrode.

Neglecting contact potential and the initial velocity of the electrons since they are of no basic importance for the understanding of this experiment, we get the following results with Hg vapor. As the potential between *K* and *G* is gradually increased up to about 4.5 volts, the current flowing through the galvanometer increases rapidly. This means that the collisions between electrons and Hg atoms are of an elastic nature. Because of their small mass, the electrons lose only a negligible part of their kinetic energy to the atoms. Evidently it is not possible in this energy range to transfer sufficient energy to an atomic electron to raise it to a higher Bohr orbit (*inelastic collision*). However, at a potential of 4.9 volts between *K* and *G* the current suddenly drops (Fig. 48), indicating that much fewer electrons arrive at *A*. This can only mean that many electrons have lost their kinetic energy by inelastic collision to the Hg atoms and so do not have sufficient energy to overcome the 0.5 volt potential difference between *G* and *A*. If the potential is increased further, the current again increases until a potential of 9.8 volts is reached. There the current again suddenly drops. At this potential the electrons can excite twice in traveling from *K* to *G*. The first inelastic collision occurs somewhere in the middle between *K* and *G*, the second one close to *G*. Figure 48 shows that a third drop of the current occurs at three times the excitation energy. The first "*critical potential*" (see page 88) of the Hg atom is thus 4.9 volts.

After a very short time this energy of the excited Hg atom is supposed to be emitted in the form of a light quantum (photon) whose frequency or wavelength is determined by Bohr's frequency condition (3-5), while the atom returns to the ground state. According to the energy relation (3-10), an excitation energy of 4.9 volts corresponds to a wave number $\bar{\nu} = 4{,}9 \times 8{,}100 = 39{,}600$ cm^{-1}, or a wavelength $\lambda = 1/\bar{\nu} = 2530$ A, which is a line in the quartz ultraviolet region. Franck and Hertz pointed a spectrograph at the space between *K* and *G* and actually found a single spectral line in the ultraviolet. The line was accurately measured to have a wavelength $\lambda = 2537$ A, in complete agreement with the computation. Thus the existence of the critical excitation potentials, predicted by the Bohr theory, was proved for the case of the Hg atom. At the same time the importance of the frequency relation (2-51) as a foundation of the Bohr theory of the atom was established in a very impressive way. *By electron impact an excitation energy of 4.9 volts was transferred to the Hg atom, and when the electron*

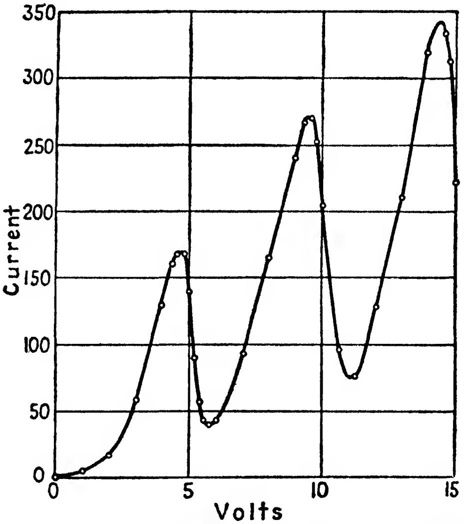


FIG. 48. Curve measured with mercury vapor by Franck and Hertz with an arrangement as in Fig. 47 establishing evidence for the first excitation potential of the Hg atom at 4.9 ev.

jumped back to the ground state, this excitation energy was emitted as a photon of the theoretically expected wavelength $\lambda = 2537$ *A.* A short time after this first experiment, many similar ones were performed

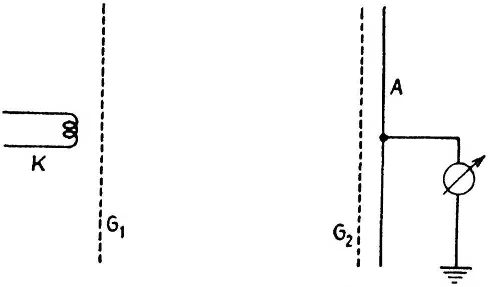


FIG. 49. Improved arrangement of Franck and Hertz for measuring higher excitation potentials. K = thermionic cathode, G_1 and G_2 = accelerating grids, A = receiving electrode.

with a large number of gases. Conditions for precisely measuring the first excitation potential are: First, the electron current density must be so small that space charge effects can be neglected. Second, the first excitation level of the atom under consideration must be well separated from the higher excitation levels. Third, the probability that the first energy level of the atom can be excited by electron collisions must be high.

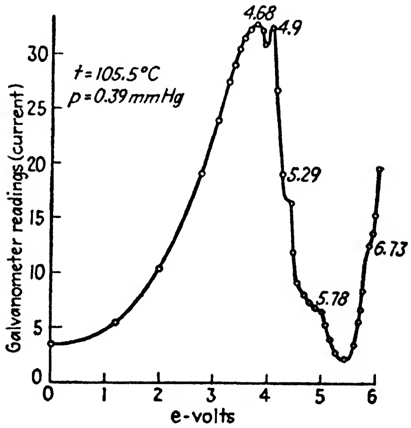


FIG. 50. Curve measured by Franck and Hertz with an arrangement as in Fig. 49, showing evidence for higher (partly metastable) energy states of the Hg atom.

By a slight modification of their experimental arrangement, Franck and Hertz succeeded in demonstrating the higher excitation levels also. The electrons were now highly accelerated in the region between K and G_1 by the grid, G_1, Fig. 49. The pressure in the chamber was greatly reduced to avoid exciting impacts in this region. Between G_1 and G_2 the electrons were gradually accelerated within that potential region where higher excitation potentials were expected. With this arrangement, Franck and Hertz obtained the curve shown in Fig. 50, and thus determined the indicated new critical potentials of the mercury atom. Among these newly found energy levels of the mercury atom are those at 4.68 and 5.29 ev. Since transitions from these levels to the ground state were not observed as spectral lines, these transitions were called "forbidden." These energy states are also called metastable (see page 139) because they have a long life and do not immediately revert to a lower state by radiating their energy. These metastable states were first discovered by this electrical method.

With the higher energy levels an increasing number of spectral lines is expected. From the relation (3-10) it can be easily computed that the transition from the 6.73-ev level to the ground state is accompanied by the emission of the shortwave Hg line 1849 A. However, as will be understood later, the transition which is expected from the combination principle, $\Delta E = 6.73 - 4.9$ ev, is forbidden.

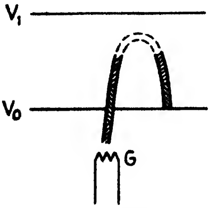


Fig. 51. Arrangement of Gehrcke and Seeliger for stepwise excitation of spectral lines by stopping a cathode ray in an opposing electric field.

With a qualitative optical method, Gehrcke and Seeliger had already observed the step-by-step excitation of the various lines of an atomic spectrum before the publication of the Bohr theory. They directed an oblique cathode-ray beam into a low-pressure gas, stopping it with an opposing field in the manner illustrated in Fig. 51. Near the vertex of the parabola the electron energy was lower than the lowest excitation energy of the atom so that no excitation could occur there, and that part of the beam was invisible. However, near this invisible part, the cathode-ray beam displayed regions of different color, caused by the gradual appearance of the lines corresponding to increasing excitation potentials. Later a similar experiment using a special glow discharge was performed

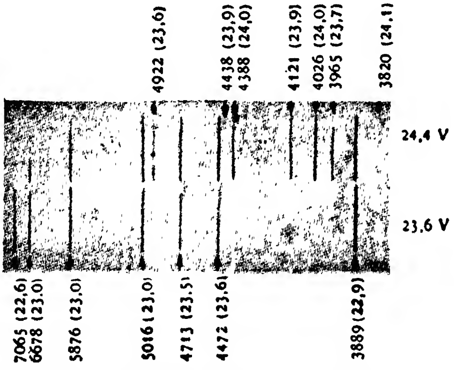


Fig. 52. Stepwise excitation of the spectrum of the Ne atom. Note the higher number of lines excited with 24.4 volts compared with 23.6 volts. (*After G. Hertz.*)

by Lau and Reichenheim and used successfully in an investigation of the H_2 spectrum, carried out in cooperation with the author. Finally, G. Hertz measured the step-by-step excitation of a spectrum very exactly. Figure 52 shows an example of his results.

The emission mechanism discussed so far is called *excitation emission*. Fundamentally different from it is the *recombination emission*, which occurs when an electron is "captured" by an ion and then emits radiation in jumping to lower states. We shall consider details of this process on page 109. Here we only point to it as a process different from excitation emission.

Up to now we have discussed only the optical and electrical evidence of excited states and the measurement of excitation energies, but we have not discussed the measurement of ionization energies. Fundamentally, in a plot like Fig. 48, we expect a sudden decrease in the current at the ionizing potential also. However, no distinction between excitation and ionization seems possible by that method. Therefore, almost all collision experiments whose purpose was to determine ionization potentials used the space-charge effect of the resulting positive ions. The appearance of positive ions can be easily detected because they reduce the negative space charge of the electrons and thus cause a sharp increase in current at the same applied potential. A clear illustration of this method is shown by the apparatus used by G. Hertz, Fig. 53. Electrons emitted by the incandescent cathode G are accelerated between G and N_2 and collected by electrode P. The electron current between G and P is small because it is hindered by the negative space charge built up between G and N_2. However, it increases rapidly as soon as the ionization potential of the gas is exceeded and the negative space charge is reduced by the positive ions. The second cathode D and the grating N_1 are used to measure the first excitation potential of the same gas, but we are not interested in that here. Figure 54 shows the ionization potential of the neon atom and, in addition, the excitation potential of Ne, determined by the difference method.

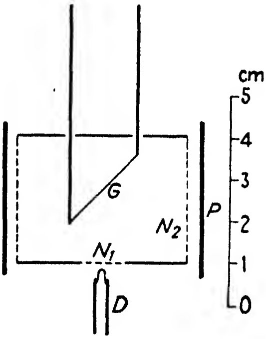


Fig. 53. Arrangement of Hertz for measuring the excitation and ionization potential of atoms.

We recapitulate: *From electron collision experiments which were carried out in various ways, Bohr's relationship between the excited states, the term values of the spectral lines, and the ionization energies was completely verified. We are thus justified in regarding these measurements as the experimental confirmation of the postulates of the Bohr theory.*

Because excitation and ionization by electron collisions are so very important in the whole field of atomic physics and in many applications such as gas discharges and light production, we shall briefly discuss the yield of these collision processes. Up to now we have only asked what electron energies are necessary to produce excitation or ionization. Now we ask, what is the probability that an electron collision of given energy actually results in an excitation or ionization, respectively? The relation between excitation probability or ionization probability and the energy of the colliding electrons in electron volts is called the excitation function or ionization function, respectively. Figures 55 and 56 present examples. The excitation probabilities can, in principle, be measured electrically by the above methods, if the number of collisions and the number of slow electrons resulting from the collisions are measured. However, the values obtained from this method are not

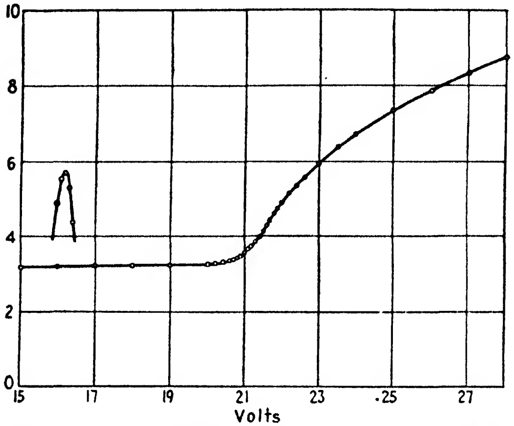


FIG. 54. Determination of the ionization potential (and the first excitation potential) of the Ne atom with the arrangement shown in Fig. 53, by Hertz.

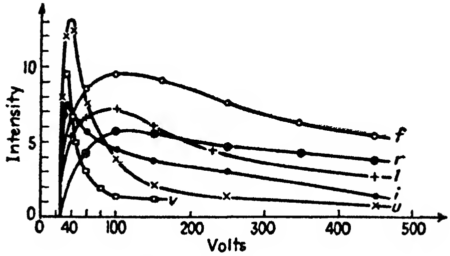


FIG. 55. Optical excitation functions showing the line intensity (proportional to the excitation probability) of some Hg lines as a function of the energy of the exciting electrons. *f*, *r*, *l* = singlet lines, *u*, *v* = triplet lines. (*After Hanle.*)

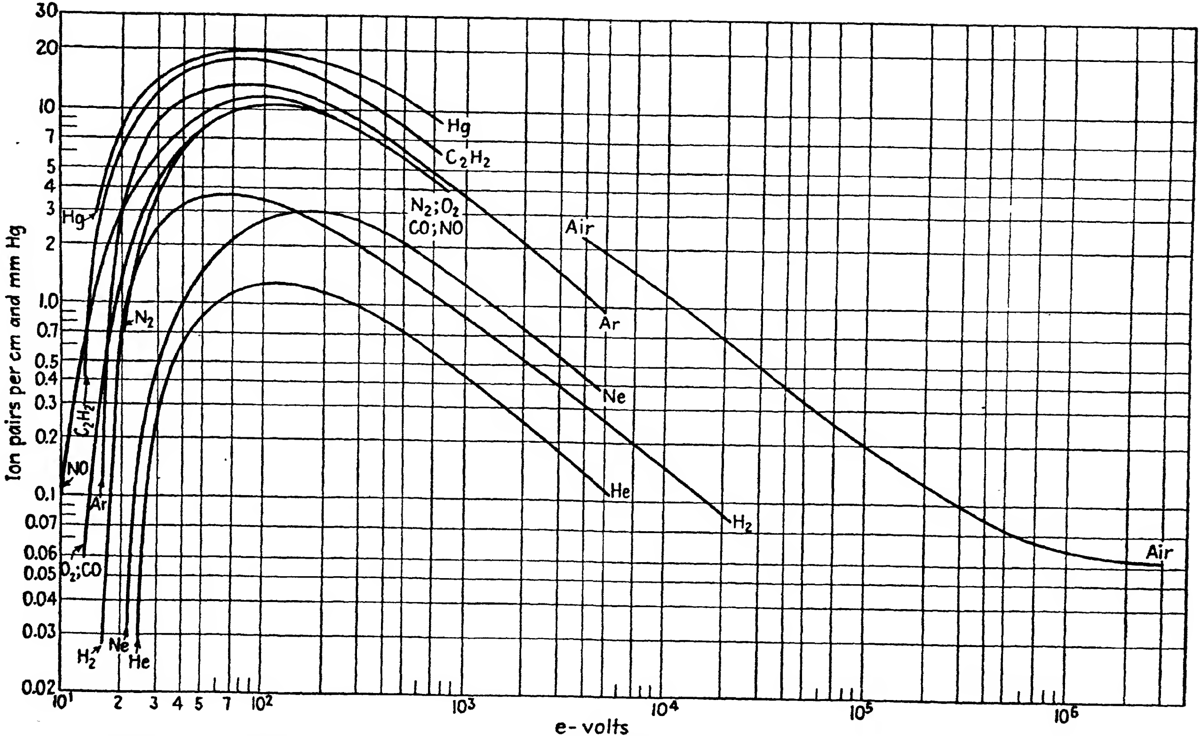


FIG. 56. Ionization functions of some atoms and molecules, showing the number of ions produced by one electron per centimeter path (proportional to the ionization probability) as a function of the energy of the ionizing electrons. (*After von Engel and Steenbeck.*)

very accurate because of several sources of error. By this method the excitation probabilities for various lines have been found to be between 1 and almost 100 per cent. More accurately known are the relative shapes of the excitation functions. To measure these relative excitation functions, the intensity of the spectral line arising from a certain excited state is plotted as a function of the energy of the exciting electrons. An example of several optical excitation functions of this type is shown in Fig. 55. Generally, the excitation probability rises steeply above the critical potential and, after passing through a maximum, decreases more or less rapidly. The position of the maximum and the slope of the curve beyond the maximum vary considerably for different types of lines. This is well illustrated in Fig. 55 for the singlet and triplet lines of a two-electron system such as the helium or mercury atom. These differences of the excitation functions will be explained theoretically on page 148.

There are several methods for determining ionization functions. Basically, in all of them the number of ions produced per centimeter of electron path is measured as a function of the energy of the colliding electron. Some results of such measurements are plotted in Fig. 56. We see that the ionization functions of different atoms and molecules have very similar shapes.

3-5. The Hydrogen Atom and Its Spectra According to the Bohr Theory

In the last section we discussed in detail the experimental proof of the Bohr postulates for the existence of discrete, stationary energy states of the atoms, which converge toward their ionization energies. We also discussed the relationship between these excited states and the emission of spectral lines as a consequence of transitions between these states, i.e., the experimental evidence for the correctness of the Bohr frequency condition (3-5). We shall now apply the Bohr theory to the simplest atom, the H atom, and its spectra. The line spectrum of the H atom in the visible and near ultraviolet region is shown in Fig. 43. This spectrum is described with high accuracy by the Balmer formula

$$\bar{\nu} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \qquad n = 3, 4, 5, \ldots \tag{3-1}$$

How does the Bohr theory explain these facts?

As the first atom in the periodic table, the H atom has the atomic number 1. Consequently it consists of only one proton for a nucleus, and, according to the Bohr concept, one electron rotating about it, in a stationary orbit. If we let the radius of this circular orbit of the electron be r, the mass of the electron be m, and its angular velocity ω, the system will be in equilibrium when the Coulomb force of attraction e^2/r^2 is equal to the centrifugal force on the electron, that is, when

$$\frac{e^2}{r^2} = mr\omega^2 \tag{3-14}$$

In order to determine the two unknown quantities r and ω, we need a second equation. This is the Bohr quantum condition, according to which the action, integrated over one complete revolution of the electron, must be equal to an integral multiple of the quantum of action h,

$$\oint p\,dq = nh \qquad n = 1, 2, 3, \ldots \tag{3-15}$$

We saw on page 85 that in its simplest form this integral is the product of the electron momentum by the circumference of its orbit, thus leading to equation (3-6). If, for the more general case, we substitute in Eq. (3-15) for the momentum p and the displacement dq of the electron in its circular orbit the following expressions,

$$\left.\begin{aligned} p &= mv = mr\omega \\ dq &= rd\varphi \end{aligned}\right\} \tag{3-16}$$

we obtain the quantum condition

$$\oint p\,dq = mr^2\omega\int_0^{2\pi} d\varphi = 2\pi mr^2\omega = nh \qquad n = 1, 2, 3, \ldots \tag{3-17}$$

which is in agreement with our earlier equation (3-6). From the equilibrium condition (3-14) and the quantum condition (3-17) we can solve for the two unknown quantities r and ω and obtain

$$r_n = \frac{h^2n^2}{4\pi^2m_ee^2} \tag{3-18}$$

$$\omega_n = \frac{8\pi^3m_ee^4}{h^3n^3} \tag{3-19}$$

Here r_n is the radius of the nth stationary orbit which, according to (3-18), is proportional to the square of the quantum number n. If we substitute in Eq. (3-18) the values for m_e, e, and h, we obtain for $n = 1$ the radius of the first Bohr orbit of the normal hydrogen atom,

$$r_1 = 0.529 \times 10^{-8} \quad \text{cm} \tag{3-20}$$

This value, computed theoretically from the Bohr postulates, is in such satisfactory agreement with the value obtained from the kinetic theory that we may regard this result as an outstanding triumph of the Bohr theory.

We now compute the energy states of the H atom belonging to the other quantum orbits. We write for the total energy, which is the sum of the kinetic and potential energies,[1]

$$E_n = \frac{1}{2}J\omega^2 - \frac{e^2}{r_n} \qquad (3\text{-}21)$$

where J is the moment of inertia of the rotating electron. By substituting Eq. (3-18) and (3-19) and remembering that $J = mr^2$, we get the quantized energy values of the H atom

$$E_n = -\frac{2\pi^2 me^4}{h^2n^2} \qquad n = 1, 2, 3, \ldots \qquad (3\text{-}22)$$

We consider now the spectral lines emitted as a consequence of transitions between these stationary energy states. Let the quantum numbers of the initial and final state be n_i and n_f, respectively. Then the expression for the wave numbers of all possible atomic lines of the H atom is, from (3-1) and (3-22),

$$\bar{\nu} = \frac{1}{hc}(E_i - E_f) = \frac{2\pi^2 me^4}{h^3c}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \qquad n_i > n_f \qquad (3\text{-}23)$$

By comparing Eq. (3-23) with the empirical Balmer formula (3-1) we see that the term number n of the Balmer formula is identical with the quantum number n of the Bohr theory. We also see that the Balmer formula follows from the general theoretical formula (3-23) by setting

$$n_f = 2 \qquad \text{and} \qquad n_i = 3, 4, 5, \ldots \qquad (3\text{-}24)$$

From the Bohr theory the value of the Rydberg constant R is

$$R = \frac{2\pi^2 me^4}{h^3c} = 109{,}737.31 \pm 0.04 \quad \text{cm}^{-1} \qquad (3\text{-}25)$$

which again is in excellent agreement with the spectroscopic value.

Our computations so far are not exact, however, since they were made for the limiting case of an infinitely large nuclear mass at rest with the electron rotating around it. Actually, both the electron and the nucleus rotate around their common center of gravity to which the moment of inertia in Eq. (3-17) and (3-21) is referred. The deviation is not large because of the large mass ratio of 1:1,837 of the electron to the nucleus. Nevertheless, this deviation makes it necessary that

---

[1] The potential energy is negative because it represents a binding energy (page 87). The potential is taken to be zero when the electron is removed to infinity.

instead of the Rydberg constant R_∞ (3-25), which is valid for infinite nuclear mass, we use the value

$$R = \frac{2\pi^2 m e^4}{ch^3(1 + m/M)} = \frac{R_\infty}{1 + m/M} \tag{3-26}$$

where m is the mass of the electron and M the mass of its nucleus. The value of the Rydberg constant R_{H} for the H atom computed from the above Eq. (3-26) is in complete agreement with the empirical spectroscopic value,

$$R_{\mathrm{H}} = 109{,}677.580 \pm 0.004 \quad \mathrm{cm}^{-1} \tag{3-27}$$

which is obtained from the Balmer formula. *This reduction of the empirical Rydberg constant, computed from the hydrogen spectrum, to the universal constants m_e, e, h, and c is another outstanding indication of the validity of the Bohr theory.* The extraordinary accuracy with which the Rydberg constant of hydrogen has been determined is further evidence for the unsurpassed accuracy of spectroscopy which is of especial importance for any quantitative check of the theory.

If we now plot the energy level diagram of the hydrogen atom, using the energy values obtained from (3-22), we again see that because of the identity of (3-22) with the empirical formula (3-3) we have arrived at a diagram which, except for a factor hc, is identical with Fig. 46. From the empirical Balmer formula and by considering Eq. (3-23), it follows that the lines of the Balmer series (Fig. 43) have as their common final state the state with quantum number 2. The Balmer series thus is emitted by transitions of the atom or its electron from the higher quantum states $n = 3, 4, 5, \ldots$ to the $n = 2$ state. These transitions are represented in the energy level diagram Fig. 57 by arrows, the direction of which represents the direction of the transition. The longer the arrow the greater is the energy and thus the wave number of the emitted spectral line, and correspondingly the smaller the wavelength, since $\tilde{\nu} = 1/\lambda$. The wavelengths of the Balmer lines of hydrogen converge, as can be seen from Fig. 57, toward a limit which corresponds to the binding energy of the electron in the $n = 2$ state. Since the energy values of the hydrogen atom are binding energies, they have been assigned negative values. They are numbered downward from the ionization limit which is ~~shown~~ dotted in Fig. 57. We discussed this point on page 87. The corresponding energy of the limit of the Balmer series is, according to Eq. (3-22), equal to $R_{\mathrm{H}}/4$. From this and the value (3-27) for R_{H} we can compute the short wavelength limit of the Balmer series. This is

$$\lambda_G = \frac{1}{R_{\mathrm{H}}/4} = 3647\ \mathrm{A} \tag{3-28}$$

which is in agreement with observed values. *$R_H/4$ thus is the binding energy of the $n = 2$ electron; or, conversely, it is the energy required to release an electron from this $n = 2$ state of the* H *atom, i.e., the ionization energy of the* H *atom excited to the energy state $n = 2$.* In the same way we can explain the physical significance of R_H. It is logical to inquire about the spectral lines of the H atom which correspond to the theoretical formula (3-23) when the values

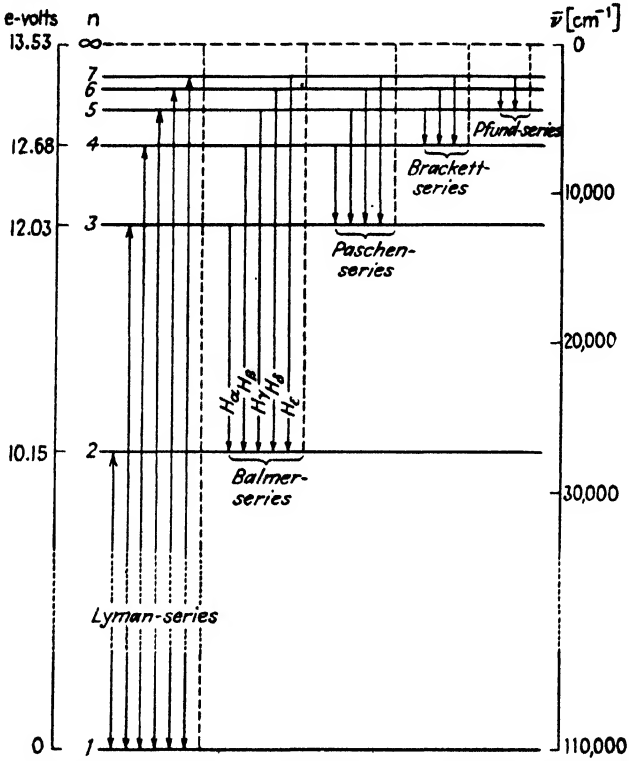


FIG. 57. Energy-level diagram of the H atom with the different spectral series.

$$n_f = 1 \qquad n_i = 2, 3, 4, \ldots \tag{3-29}$$

are used. This spectral series

$$\bar{\nu} = R_H\left(1 - \frac{1}{n^2}\right) \qquad n = 2, 3, 4, \ldots \tag{3-30}$$

corresponds to the transitions from the various excited states, enumerated on the left in Fig. 57, to the ground state, $n = 1$, of the hydrogen atom. We can compute the wavelength of the longest line of this series, the transition 2 → 1. From (3-30) it follows that for $n = 2$

$$\bar{\nu} = \frac{3}{4}R_H \qquad \lambda = \frac{1}{\bar{\nu}} = 1215.7\ \text{A} \tag{3-31}$$

The hydrogen series represented by (3-30) thus must lie in the vacuum ultraviolet and it was there discovered by Lyman. The wave number ν_G of the series limit of the Lyman series, whose terms are designated by Ly_α, Ly_β, . . . , Fig. 57, is equal to the Rydberg constant R_H. Consequently its wavelength is

$$\lambda_G = \frac{1}{R_H} = 911\ \text{A} \tag{3-32}$$

The limit lies in the far ultraviolet, where it actually has been found.

Thus the physical significance of the Rydberg constant has become clear. R_H is, as may be seen from Fig. 57, the distance of the ground state $n = 1$ from the ionization limit, and is thus the binding energy of the hydrogen electron in the normal state $n = 1$, or, conversely, the ionization energy of the atom, i.e., the energy necessary to completely separate the electron from the ground state. This important atomic constant whose experimental determination by electron impact experiments was treated in the last section thus can be determined spectroscopically with much greater accuracy. *The ionization energy is equal to the limit of that spectral series of the atom which has the shortest wavelengths.* The ionization energy of the H atom determined from the energy relation (3-10) and the R_H value (3-27) is in agreement with the value given by electron impact experiments, namely 13.595 ev. There exist also some higher spectral series of the hydrogen atom which have as end states the energy levels $n = 3$, $n = 4$, and $n = 5$ and are plotted at the right in Fig. 57. These series can also be computed from formula (3-23). They are named for their discoverers as the Paschen, Brackett, and Pfund series. Although the higher lines of the Paschen series lie in the photographic infrared, an estimate shows that the other two series lie entirely in the infrared region which is less accessible.

Thus we know the following spectral series of the hydrogen atom:

$$\left.\begin{array}{lll}
(1) \quad \bar{\nu} = R_H\left(1 - \frac{1}{n^2}\right) & n = 2, 3, 4, \ldots & \text{(Lyman series)} \\
(2) \quad \bar{\nu} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right) & n = 3, 4, 5, \ldots & \text{(Balmer series)} \\
(3) \quad \bar{\nu} = R_H\left(\frac{1}{3^2} - \frac{1}{n^2}\right) & n = 4, 5, 6, \ldots & \text{(Paschen series)} \\
(4) \quad \bar{\nu} = R_H\left(\frac{1}{4^2} - \frac{1}{n^2}\right) & n = 5, 6, 7, \ldots & \text{(Brackett series)} \\
(5) \quad \bar{\nu} = R_H\left(\frac{1}{5^2} - \frac{1}{n^2}\right) & n = 6, 7, 8, \ldots & \text{(Pfund series)}
\end{array}\right\} \tag{3-33}$$

All these series are emitted by hydrogen atoms, for example, in a glow discharge in which the H_2 molecules are dissociated into H atoms by electron collision, and then excited to the various states.

According to the Bohr theory the H atom should also be able to absorb those spectral lines which it emits. The energy corresponding to a line is absorbed by the atom, and excites it to the corresponding higher state. In Fig. 57 the transitions corresponding to absorption are represented by an arrow pointing upward. Is it now actually possible to observe in absorption all lines of the five hydrogen series (3-33)?

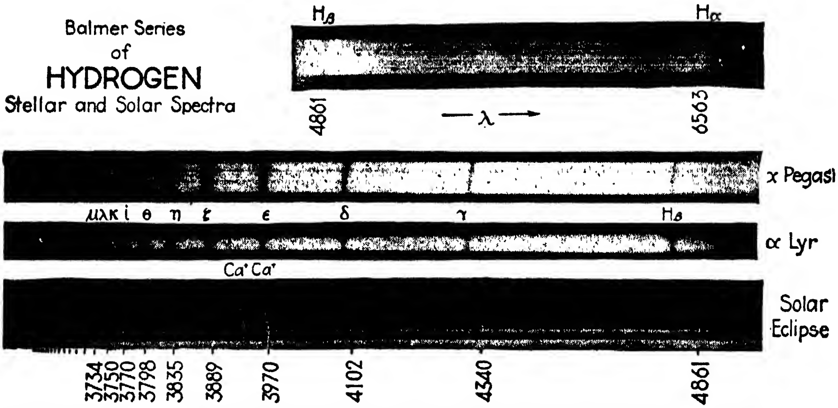


FIG. 58. Photographs of the Balmer series of the hydrogen atom emitted by the solar chromosphere (with photosphere covered by the moon) and absorbed by the atmospheres of two fixed stars. (*Courtesy of H. E. White.*)

We shall neglect for the moment the fact that hydrogen is normally a molecular gas and we shall think of it as atomic hydrogen. Then at normal temperatures there exist only normal atoms in their ground state $n = 1$. From Fig. 57 we see that the normal atom can absorb only those spectral lines whose lower state is the ground state itself; that is, for the hydrogen atom, only the Lyman series. These lines, terminating on the ground state of the atom, are called resonance lines for historical reasons. *Only resonance lines or series appear as absorption lines at normal temperatures.* Their arrows can, therefore, be shown double-headed. All higher series normally appear only in emission. This important result seems to be in disagreement with the fact that in the spectra of many fixed stars the Balmer lines, though *not* ending on the ground state, *are* observed in absorption spectra (see Fig. 58). The answer to this apparent contradiction lies in the fact that the absorbing hydrogen in the stars' atmosphere is by no means at a normal low temperature, but has a temperature of several thousand degrees, as shown clearly by the dissociation of the H_2 molecules into H atoms. For these high temperatures it follows from Maxwell's velocity distribution that a certain computable percentage of the hydrogen atoms is always excited due to thermal collisions; that is, they are in the first excited state ($n = 2$) and thus will absorb the Balmer series. Whereas at normal temperatures only resonance lines are absorbed, at sufficiently high temperatures higher series or lines can also be absorbed.

To recapitulate: *Not only are the wavelengths of the long-known Balmer lines and the value of the ionization energy of the hydrogen atom as given by the Bohr theory in excellent quantitative agreement with experimental results but, also, the theoretically predicted existence of a large number of additional spectral series has in the meantime been verified. Experiments and the Bohr theory not only agree quantitatively within the high spectroscopic accuracy, but also the relation of the various series with respect to their appearance in emission and absorption has been completely explained by the theory.*

3-6. Atomic Processes and Their Reversals. Ionization and Recombination. Continuous Atomic Spectra and Their Significance

In the last section we became acquainted with a number of important atomic processes such as excitation by electron impact and the subsequent emission of radiation, and ionization by electron impact. We shall now write these atomic processes in a somewhat different form, i.e., as reaction equations, and at the same time broaden and enhance our knowledge of these processes and the corresponding spectroscopic phenomena.

a. Collisions of the First and Second Kinds and the Resulting Processes. Emission and Absorption

We begin with the process in which an atom A is excited by electron impact, and after excitation the electron jumps back to the ground state under emission of a photon whose energy is equal to the excitation energy. We shall consider only the simple case of *one* state. If the process of exciting an atom A by electron impact is written in the form of a chemical equation, we have

$$A + \mathrm{el}_{\mathrm{fast}} = A^* + \mathrm{el}_{\mathrm{slow}} \tag{3-34}$$

where A^* is the excited atom. The fast electron gives up the excitation energy to the atom in an inelastic collision. After the collision, we have an excited atom and an electron with correspondingly smaller kinetic energy. As a result of the excitation, the atom A^* spontaneously emits a quantum (photon) of energy $hc\bar{\nu}$ and returns to the ground state. Consequently, this process can be written

$$A^* \rightarrow A + hc\bar{\nu} \tag{3-35}$$

Now there is a well-established rule in physics that every process can be reversed[1] and so we ask whether our equations (3-34) and (3-35) can be reversed and what this means. The most obvious reversal is that of (3-35). The atom A can "collide" with a photon, absorb it, and be converted into the excited atom A^*. *By reading Eq. (3-35) from left to right we have emission, and from right to left, absorption, of radiation.*

The reversal of the excitation by electron collision (3-34) means that an excited atom collides with an electron and gives up its excitation energy in the collision, so that we get a normal atom and an electron of correspondingly higher kinetic energy. This process is called a "collision of the second kind" as compared with excitation by electron collision which is called a "collision of the first kind." Collisions of the second kind, in which atoms or molecules give up their excitation energy without radiation, do not require collisions with electrons only, but they can occur with other particles. For instance, an excited atom A^* can collide with another atom or molecule B which has a possible excited state of the same energy as the excitation energy of A^*. In a collision of the second kind, A^* can then give up its energy to B. Then we have the process

$$A^* + B \rightarrow A + B^* \tag{3-36}$$

In contrast to reversal of (3-34)

$$A^* + \mathrm{el}_{\mathrm{slow}} \rightarrow A + \mathrm{el}_{\mathrm{fast}} \tag{3-37}$$

the process (3-36) can occur only if A and B have approximately the same excitation energies or if, in other words, there is resonance between the energy A^* and one of the many energy states of B.

The radiation from atoms or molecules B as a result of collisions of the second kind with excited atoms according to (3-36) is called "sensitized fluorescence" after Franck. It can be easily observed in a mixture of different metal vapors with mercury vapor. The collisions of the second kind with electrons, according to (3-37), studied first by Klein and Rosseland, occur with great probability (yield) whenever collisions between excited atoms or molecules and electrons occur. It is, of course, necessary for both collisions of the second kind (3-36)

---

[1] It is obvious that for the reversal of certain processes a large expenditure of energy is required.

and (3-37) that the collision occurs before the excited atom releases its energy in the form of radiation. Since, generally, the emission of radiation occurs after the mean life-time of the excited state of about 10^{-8} sec, it can be shown from the kinetic theory of gases how the probabilities of the concurrent processes of radiation and collision of the second kind depend on the gas pressure. We have already mentioned on page 92 and shall treat in detail, especially in the next section, cases in which the excited atom or molecule cannot radiate. Such excited states are called *metastable* and the corresponding particle is called a metastable atom or molecule. Collisions of the second kind therefore occur predominantly in the presence of atoms or molecules with metastable states, such as the mercury atom.

b. Impact Ionization and Recombination in Triple Collisions

We now turn our attention from exciting collisions and their reverse processes to the ionization by electron collisions wherein an electron with sufficiently high kinetic energy collides with an atom and ionizes it. After the collision we have a positive ion A^+, an electron resulting from the ionization, and the colliding electron with correspondingly diminished energy. The process of impact ionization by an electron thus can be written:

$$A + \mathrm{el}_{\mathrm{fast}} \rightarrow A^+ + \mathrm{el} + \mathrm{el}_{\mathrm{slow}} \tag{3-38}$$

This process can also be reversed, i.e., we can read the above equation from right to left: A positive ion A^+ undergoes a triple collision with two electrons and recombines with one of them to form a neutral atom, which sometimes is first excited and then reverts to the ground state by emission. The second electron in this triple collision carries away as kinetic energy the binding energy of the recombined atom, which is equal to the ionization energy. This *triple-collision recombination* of an ion and electron into a neutral atom is one of the most important processes in all electric gas discharges, in which charge carriers are eliminated mainly by this process.

Since in this reversed process (3-38)

$$A^+ + \mathrm{el} + \mathrm{el} \rightarrow A + \mathrm{el}_{\mathrm{fast}} \tag{3-39}$$

the second electron has no other task than to carry away the binding energy when the ion recombines with the other electron, it can be replaced by any other particle B, an atom, molecule, or ion. Recombination in triple collisions, therefore, instead of (3-39), can also occur in the similar process

$$A^+ + \mathrm{el} + B \rightarrow A + B_{\mathrm{fast}} \tag{3-40}$$

Instead of colliding with a third particle *B*, the recombining ion and electron may finally encounter any larger body such as the wall of a container and release their binding energy there, thus heating up the body. Recombination of ions and electrons occurs at the wall of almost all discharge tubes and accounts, at least partly, for their high temperatures.

In the processes (3-39) and (3-40) the recombination occurs in triple collisions because the binding energy is carried away by the third particle. A recombination in a two-particle collision is possible if the binding energy is carried away without the aid of a third particle, that is, if it is emitted as radiation. This process is also possible, and its discussion leads to a much more complete understanding of the relation between spectra and atomic processes.

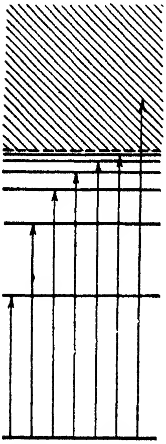


FIG. 59. Discrete energy states and continuous energy region of an atom, corresponding to excitation and ionization of the atom, respectively. Arrows indicate transitions corresponding to the line spectrum and the series limit continuum.

c. Photoionization and the Series Limit Continuum in Absorption

We have already discussed photoexcitation, the excitation of electrons by the absorption of energy (i.e., a photon), according to

$$A + hc\bar{\nu} \rightarrow A^* \tag{3-41}$$

We also know that the series of excitation states of an atom converges toward the ionization energy, so that the idea suggests itself that ionization of an atom should also be possible by absorption of the necessary energy in the form of a photon. In this case, which actually occurs, we speak of photoionization. In order to understand the process correctly, we must expand the energy level diagram of the atom, Fig. 59. So far we have discussed only the stationary energy states of the atom. The higher the energy state, the weaker is the binding between the nucleus and the electron. The broken line represents the limit of convergence of the energy states and corresponds to a free electron completely removed from its atom. What happens now if we impart more energy to the ground state than is necessary to ionize the atom? Obviously the electron will be released and will take with it the excess energy as kinetic energy, $mv^2/2$. By extending our usual concept of the atom, we can also regard this state of the system, ion + electron with kinetic energy, as a state of the atom. We speak then not of a ***stationary state of the bound electron***, but of a ***nonstationary state*** *of the free electron.* Since the energy of the electron is computed from the ground state of the atom, and its energy is greater than the ionization energy, its state must lie above the ionization limit. Naturally these free states are not quantized since the free electron can have any arbitrary amount of kinetic energy. Our energy level diagram of the atom, Fig. 59, thus has, in addition to the discrete energy levels, a continuous energy range of nonstationary states of the free electron above the ionization limit.

What is the result of this discussion of the absorption spectrum of the atom? The transitions from the ground state to the different excited states of the atom, plotted as arrows in Fig. 59, correspond to the absorption process (3-41). They produce a line series (for the H atom, the Lyman series) which converges toward the ionization limit. But the spectrum does not end here. Beyond the series limit, extending toward the shorter wavelengths, there is a continuous spectrum (see the absorption spectrum of the Na atom, Fig. 41) which is a result of transitions from the ground states to the free states lying above the ionization limit. That the absorption intensity or absorption coefficient of this continuum beyond the series limit falls off toward shorter waves, follows from intensity rules to be discussed later. This means that the probability of absorption for large energy photons is smaller than for those energies which exceed the ionization energy by only small amounts.

This process, the ionization of an atom by absorption of a photon with a wave number corresponding to the series limit continuum, is called photoionization. It may be written

$$A + hc\bar{\nu}_c \rightarrow A^+ + \text{el} + \frac{m}{2}v^2 \tag{3-42}$$

From observations of this series limit continuum in absorption we conclude that photoionization is an important atomic process. Quantitatively, the intensity ratio of the atomic lines to the series limit continuum gives us the relative probability of photoexcitation and photoionization processes. The intensity distribution of the series limit continuum, on the other hand, gives us the probability of photoionization as a function of the energy $hc\bar{\nu}$ of the absorbed photon.

d. Radiative Recombination and the Series Limit Continuum in Emission

The process of photoionization (3-42) can be reversed to give us the process

$$A^+ + \text{el} + \frac{m}{2}v^2 \rightarrow A + hc\bar{\nu}_c \tag{3-43}$$

A positive ion A^+ collides with an electron and recombines to form an atom A. In most cases, A is first excited and then reverts to the ground state by line emission. The sum of the atomic binding energy (= ionization energy) and the relative kinetic energy of ion and electron is emitted as a photon with a wave number in the continuum. In the language of the energy level diagram this means that, in addition to the transitions from the higher stationary states to the ground state, there are transitions from the continuous energy region above the ionization limit to the ground state. As a consequence of these transitions, the continuum is emitted.

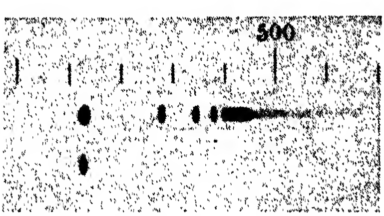


Fig. 60. Emission lines and series limit continuum of the principal series of the He atom in the extreme ultraviolet. (*Photograph by Lyman.*)

From the ionization energy of 24.5 ev measured for the He atom, it is possible to compute, using (3-10), the wave number of the limit of the resonance series to be about 200,000 cm^{-1} and from it the wavelength to be about 500 A. The continuum emitted according to (3-43) during the radiative recombination of the He^+ ion with an electron into the ground state of the normal atom, thus should extend from this wavelength to shorter waves. This continuum, as shown in Fig. 60, was actually found in a helium discharge by Lyman with his vacuum spectrograph. Its long wave limit, 509 A, is in good agreement with our estimate. *We conclude from the observation of emission continua that the process of recombination of atoms and electrons in two-particle collisions with the emission of the binding energy according to (3-43) actually occurs. From the intensity distribution in the continua we can draw conclusions about the frequency of the recombination of ions with electrons of various kinetic energies.* The decrease of intensity toward shorter waves indicates that radiative recombination occurs predominantly with slow, and much more rarely with fast, electrons.

Observations on the emission and absorption continua and measurements of their intensity distribution are important for the astrophysicist because from these it is possible to draw conclusions about the atomic processes and the state of the radiating or absorbing layers of the stars. It is fortunate in this respect that the radiative recombination (3-43) is more frequent than the triple-collision recombination (3-39) or (3-40) at the low gas pressures of many astronomical objects. The high intensity of the continuous spectra of a discharge in thallium vapor, Fig. 61, illustrates that the radiative recombination (3-43) is a very important process in many discharges also, in spite of the fact that at high pressures the triple-collision recombination (3-39) or (3-40) usually predominates.

The radiative recombination need not necessarily lead immediately to the ground state, but the capture of an electron often occurs in an excited state, and then is followed by the emission of a spectral line when the electron jumps to the ground state. We have already mentioned this process of line emission under the name of recombination radiation on page 93. Since the process of transition from the higher, excited H states to the second quantum state, Fig. 57, corresponds to the emission of the Balmer series, we have as a result of the transition of free electrons to the second quantum state, the emission of the Balmer continuum beyond the shortwave limit of the Balmer series. Similarly, the continuum beyond the shortwave limit of the Paschen series is emitted upon the capture of electrons with various kinetic energies into the third quantum state of the H atom. This spectrum has been observed in a condenser discharge in hydrogen. To make such observations it is best to use low-pressure discharges (in order to suppress triple-collision recombination) with reasonably high current densities. For a high current density means a sufficiently high density of electrons and ions and thus a large number of collisions. Figure 61 shows the spectrum of such a discharge in thallium vapor in which the emission continua of the various series are clearly distinguishable.

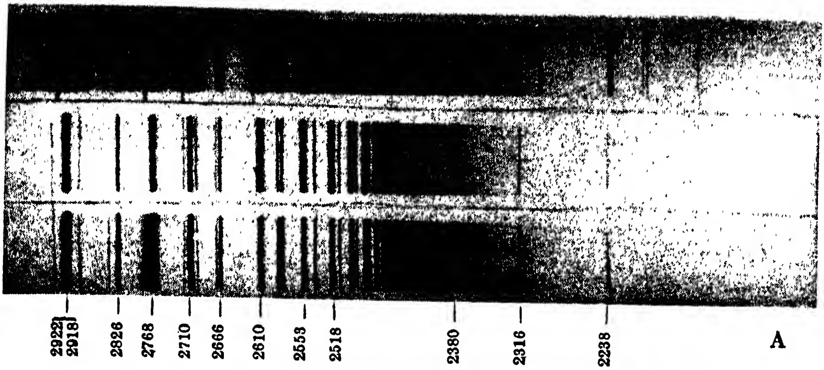


Fig. 61. Series limit continuum of the subsidiary series of the Tl atom emitted by a thallium vapor discharge. With increasing current density (from the lower to the upper spectrum) the continuum penetrates more and more across the series limit into the series, caused by the disturbing fields of electrons and ions. Simultaneously forbidden lines appear. (*Photograph by courtesy of H. Krefft.*)

e. Electron Bremsstrahlung

The continua which we have just discussed are represented in the energy level diagram as transitions from the continuous energy region to the different stationary states of the atom. For the sake of completeness of our theoretical picture, it is finally expected that there are also transitions between free electron states within the continuous energy region, Fig. 62, corresponding to either absorption or emission of radiation. Since the initial and final states now lie in a continuous energy region, the resulting spectrum, even though it is continuous, should have no direct relation with any line spectra. To what process do the transitions shown in Fig. 62 correspond? If we consider the transition from higher to lower states, then in the initial state the free electron has a higher kinetic energy than it has (after the transition) in the lower state. The energy difference is emitted as a photon with a wave number of the continuous spectrum. We thus are dealing with the *process of stopping a fast electron in the field of a positive ion.* However, in contrast to the recombination, the electron in this case is not captured into a stationary state but, after collision with the ion, continues on, though with reduced kinetic energy. This process can be written as

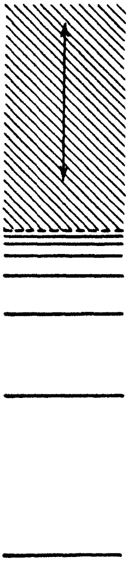


FIG. 62. Transitions corresponding to the electron retardation continuum represented in the energy-level diagram of an atom.

$$A^+ + \mathrm{el}_{\mathrm{fast}} \rightarrow A^+ + \mathrm{el}_{\mathrm{slow}} + hc\bar{\nu}_c \qquad (3\text{-}44)$$

A continuous spectrum resulting from this process is called an *electron retardation continuum,* or, by the untranslatable German expression, an *electron bremsspectrum.* Its shortwave limit depends upon the kinetic energy of the fastest colliding electrons. If this is $(m/2)v^2$, certainly not more than the total kinetic energy can be radiated in the stopping process (3-44) so that the wave number $\bar{\nu}_G$ of the shortwave limit is given by the relation

$$hc\bar{\nu}_G = \frac{m}{2}v^2 \qquad (3\text{-}45)$$

From this limit the continuum can extend without limit toward longer wavelengths, since arbitrarily small energy changes may occur in the retardation process.

As in all atomic processes we have treated so far, the process of retardation radiation (bremsstrahlung) can be reversed. Reading

Eq. (3-44) from right to left means that an electron colliding with an ion absorbs continuous radiation and gains kinetic energy. Then the transition arrow in Fig. 62 is directed upward. This last process plays an important role in astrophysics.

An electron retardation continuum (3-44) was first found when fast cathode-ray electrons were stopped by colliding with Z-fold charged nuclei of the metallic atoms in the anticathode of an X-ray tube. For example, if an electron of 50,000 volts (50 kv is a normal potential for an X-ray tube) is stopped in the nuclear field of an anticathode atom, then from the energy relation (3-10) the wave number is about 4×10^8 cm^{-1} and the corresponding wave length $\lambda = 1/\bar{\nu}$ is 2.5×10^{-9} cm = 0.25 A, i.e., belongs to the X-rays (Fig. 24). The long-known continuous X-ray spectrum thus is due to the process (3-44). This spectrum is called the *X-ray retardation continuum.* In our illustration in Fig. 62, since the ionization energy of the atom is only of the order of magnitude of 10 ev, the upper end of the arrow would lie 50,000 ev or 250 m (!) above the ground state. Until recently it was not definitely proved that electron retardation continua occur also in the optical region. However, recently the author has been able to show that a large number of continuous spectra, especially of arc and spark discharges of high current density, are definitely retardation continua. Retardation radiation in the visible region is also of importance in astrophysics.

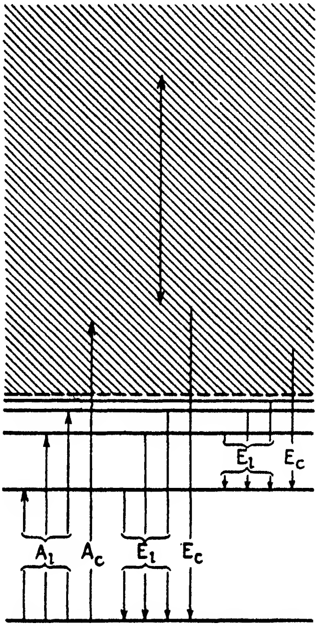


FIG. 63. Schematic representation of all possible transitions in the energy-level diagram of an atom. A_l = absorption of spectral lines; A_c = absorption of series limit continuum; E_l = emission of spectral lines; E_c = emission of series limit continua; uppermost double arrow = electron retardation continuum (emission or absorption, respectively).

In Fig. 63 an energy level diagram is shown with all possible atomic transitions corresponding to emission as well as absorption of lines, series limits continua, and retardation continua (bremscontinua).

3-7. The Spectra of Hydrogen-like Ions and the Spectroscopic Displacement Law

Before the publication of the Bohr theory, in addition to the spectral series (3-33), two further series were ascribed to the hydrogen atom, which are similar in their structure to the Balmer series. One of these series was found by Fowler in a hydrogen-helium discharge, the other by Pickering in the spectra of certain fixed stars. The Balmer series of hydrogen, as measured by Paschen, and the questionable Pickering series are shown in Fig. 64. Each second line of this series thus falls close to a Balmer line, whereas the other lines fall in between the Balmer lines.

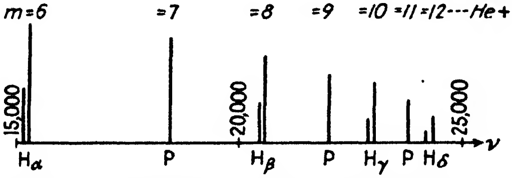


FIG. 64. Comparison of the wavelengths of the Balmer series of the H atom and the Pickering series (indicated by P) of the He$^+$ ion. (*After Sommerfeld.*)

The two questionable series can be described by the following Rydberg formulas:

$$\text{Fowler series:} \quad \bar{\nu} = 4R\left(\frac{1}{3^2} - \frac{1}{n^2}\right) \qquad n = 4, 5, 6, \ldots \tag{3-46}$$

$$\text{Pickering series:} \quad \bar{\nu} = 4R\left(\frac{1}{4^2} - \frac{1}{n^2}\right) \qquad n = 5, 6, 7, \ldots \tag{3-47}$$

How can these formulas be explained by the Bohr theory? The H atom definitely cannot be responsible for the spectra since from our computations on page 101 there is no room in the theory for series of this type. Theoretically, the spectrum of the hydrogen atom is complete with the series (3-33). Since the new series were observed by Paschen in pure helium also, the idea suggests itself to ascribe them to a singly positively charged helium ion, He$^+$. This explanation has been verified and led to a number of important conclusions.

We say that the He$^+$ ion is hydrogen-like in a rigorous sense because, like the H atom, it has a nucleus and one orbital electron. Because of the double charge on the helium nucleus we use in the Bohr theory of the He$^+$ ion, instead of the equilibrium condition (3-14), the condition

$$\frac{2e^2}{r^2} = mr\omega^2 \tag{3-48}$$

The further theoretical treatment is analogous to that of Sec. 3-5, page 97. There we obtained the energy values of the hydrogen atom

$$E_n(\mathrm{H}) = -hc\frac{R}{n^2} \qquad n = 1, 2, 3, \ldots \tag{3-49}$$

Because of the doubly charged nucleus of He$^+$ we now get the energy values

$$E_n(\mathrm{He}^+) = hc\frac{4R}{n^2} \qquad n = 1, 2, 3, \ldots \tag{3-50}$$

Quite generally, we find for a nucleus with charge $+Z$

$$E_n(Z) = -hc\frac{Z^2R}{n^2} \qquad n = 1, 2, 3, \ldots \tag{3-51}$$

The energy level diagram of the He$^+$ ion thus is exactly the same as that of the H atom except that the energy differences are increased by a factor 4 (or in general Z^2). The ionization energy necessary to remove the remaining electron of the He$^+$ ion, therefore, is $4R = 54.4$ ev. The wave number $\bar{\nu}$ of the spectral lines of He$^+$ is given by (3-51) and is

$$\bar{\nu} = 4R\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \qquad n_i > n_f \tag{3-52}$$

We thus have again the generalized Balmer formula (3-23) of page 98 except that the factor R is replaced by $4R$. All lines of He$^+$ have only one fourth the wavelength of the corresponding lines of the H atom. This means that they are shifted into the ultra-violet. A comparison of (3-52) with (3-46) and (3-47) shows that the Fowler series and the Pickering series are clearly series of He$^+$. Why each second line of the Pickering series coincides with a Balmer line can be understood if the formula for the Pickering series (3-47) is written

$$4R\left(\frac{1}{4^2} - \frac{1}{n^2}\right) = R\left[\frac{1}{2^2} - \frac{1}{(n/2)^2}\right] \qquad n = 5, 6, 7, \ldots \tag{3-53}$$

For even n, this gives those lines of the Pickering series which coincide with Balmer lines, for odd n those between the H lines.

It has already been pointed out in connection with Fig. 64 that in contrast to our computations (3-53) lines with even n do not exactly coincide with Balmer lines either. Table 6 gives the approximate wavelengths of the first lines of both series from measurements of Paschen. The wavelength differences can be taken from them. This noncoincidence depends on the difference between the Rydberg constants R for hydrogen and helium, which in turn is due to the different effect of the nuclei of H and He on the revolution of nucleus and electrons around their common center of mass. The Rydberg constant R_∞ was deduced on page 98 for an infinitely large nuclear mass M, Eq. (3-25). By taking into account the actual motion of the nucleus and electron around their common center of mass, this value was reduced in the ratio

$$1 + \frac{m}{M} : 1 \tag{3-54}$$

Table 6. Comparison of the Wavelengths of the Lines of the Pickering Series of He^+ *and the Balmer Series of the* H *Atom*

Upper quantum number of Pickering series	He^+	H
6	6560.1	6563.1 (H_α)
7	5411.6	
8	4859.3	4861.3 (H_β)
9	4561.6	
10	4338.7	4340.6 (H_γ)
11	4199.9	
12	4100.0	4101.9 (H_δ)

Consequently, since the mass of the helium nucleus is four times that of the hydrogen nucleus, the Rydberg constant of He^+ is

$$R_{\text{He}} = \frac{R_\infty}{1 + (m_e/M_{\text{He}})} = \frac{R_\infty}{1 + (m_e/4M_{\text{H}})} \tag{3-55}$$

From the spectra of He^+, i.e., from the Rydberg formulas (3-46) and (3-47), the value of R_{He} has been found to be in agreement with (3-55). Instead of the R_{H} value (3-27), we have

$$R_{\text{He}} = 109{,}722.269 \pm 0.00024 \text{ cm}^{-1} \tag{3-56}$$

From the agreement between computed and spectroscopic results the important conclusion can be drawn that the law of center of mass, which was used to compute the motion of the nucleus, is also valid in atomic systems.

According to Bohr, from (3-55) and the corresponding equation for R_{H}, an important universal constant, the mass ratio of proton and electron, can be computed:

$$\frac{M_{\text{H}}}{m_e} = \frac{R_{\text{H}} - \frac{1}{4}R_{\text{He}}}{R_{\text{He}} - R_{\text{H}}} \tag{3-57}$$

Since spectroscopists can measure R_{H} and R_{He} with a high degree of accuracy, the mass ratio (3-57) can be determined with equal accuracy. The present-day value is

$$\frac{M_p}{m_e} = 1{,}836.6 \pm 0.2 \tag{3-58}$$

Now the specific charge of hydrogen ions e/M_{H} can be determined with high accuracy from electrolysis (Faraday's law). Its value is the so-called electrochemical equivalent of 96,496 coulombs/gm. From this and (3-58) we get a semispectroscopic value for the specific charge of the electron e/m_e (see page 35).

After this discussion of the significance of the Rydberg constant we return to the spectral series of He^+. Of the series predicted by (3-52), we have already discussed the Fowler series ($m_f = 3$) and the Pickering series ($n_f = 4$). The two shortwave series with $n_f = 2$ and $n_f = 1$ have been found in the shortwave vacuum ultraviolet and are in quantitative agreement with (3-52) as was expected by Lyman. Consequently, analogous to the five series (3-33) of the H atom, we have the following series for the hydrogen-like He⁺ ion:

$$\begin{array}{lll}
\text{First Lyman series:} & \bar{\nu} = 4R_{\text{He}}\left(1 - \dfrac{1}{n^2}\right) & n = 2, 3, 4, \ldots \\[2ex]
\text{Second Lyman series:} & \bar{\nu} = 4R_{\text{He}}\left(\dfrac{1}{2^2} - \dfrac{1}{n^2}\right) & n = 3, 4, 5, \ldots \\[2ex]
\text{Fowler series:} & \bar{\nu} = 4R_{\text{He}}\left(\dfrac{1}{3^2} - \dfrac{1}{n^2}\right) & n = 4, 5, 6, \ldots \\[2ex]
\text{Pickering series:} & \bar{\nu} = 4R_{\text{He}}\left(\dfrac{1}{4^2} - \dfrac{1}{n^2}\right) & n = 5, 6, 7, \ldots
\end{array} \qquad (3\text{-}59)$$

The next hydrogen-like ion after He^+ is the Li^{++} ion which consists of a triply charged positive nucleus and one orbital electron. Its spectra are given by the Rydberg formula

$$\bar{\nu} = 9R\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \qquad n_i > n_f \qquad (3\text{-}60)$$

in which the Rydberg constant is determined by substituting the mass of the lithium nucleus in Eq. (3-26). Unfortunately, only the first three terms of the resonance series of the Li^{++} ion have been observed, since they lie in the extreme vacuum ultraviolet. The shortest wave line that has been observed by normal ultraviolet technique is the spectral line of the elevenfold ionized aluminum atom at 6.31 A. This ion, however, is not hydrogen-like since it still has two electrons.

In general, the energy value of a hydrogen-like ion with a charge Z on the nucleus and one electron is given by

$$E_n = \frac{hcZ^2R}{n^2} \qquad n = 1, 2, 3, \ldots \qquad (3\text{-}61)$$

If all electrons but one could be stripped from an atom, for example from the calcium atom with $Z = 20$, the ionization energy of this ion would be $20^2R = 400R = 4.4 \times 10^7\ \text{cm}^{-1}$. The value for the wavelength of the series limit would be $\lambda_G = 1/400R = 2.3$ A. This wavelength is beyond the optical region, in the X-ray region. We see that X-rays are emitted by transitions of the innermost Ca electron in the strong Coulomb field of the Ca nucleus, and we see that there is a continuous transition from the optical to the X-ray region. We will discuss this in further detail in the next section.

So far we have discussed only those ions which are hydrogen-like in the most precise sense. The spectra of all other atoms and ions with more than one electron have rather complicated structures. The electrons of these atoms and ions are arranged in shells, and one can speak of hydrogen-like ions in a wider sense in the case of the alkali atoms and of ions which have only one electron in their outer shell, which upon excitation is responsible for the spectra. Energy states of these hydrogen-like atoms and ions can no longer be computed according to Eq. (3-61) because the electrostatic binding between the outermost "valence electron" and the nucleus is disturbed by the presence of the other electrons. We shall consider the details of the spectrum of the alkali atoms on page 125.

We saw that the spectra of the sequence H, He^+, Li^{++} were identical except for the shift towards the ultraviolet, caused by the increasing nuclear charge. We expect the same behavior for the spectra of the following ions, which are still hydrogen-like to some approximation:

$$\mathrm{Li},\ \mathrm{Be}^{+},\ \mathrm{B}^{++},\ \mathrm{C}^{+++},\ \mathrm{N}^{++++},\ \mathrm{O}^{5+},\ \mathrm{F}^{6+},\ \mathrm{Ne}^{7+}$$
$$\mathrm{Na},\ \mathrm{Mg}^{+},\ \mathrm{Al}^{++},\ \mathrm{Si}^{+++},\ \mathrm{P}^{++++},\ \mathrm{S}^{5+},\ \mathrm{Cl}^{6+}$$

All these ions, like the alkali atoms lithium and sodium, have one electron in their outer shell. Their spectra and energy-level diagrams therefore should be equal except for an increasing shift toward the ultraviolet. This has been confirmed by spectroscopic investigations. Millikan and Bowen, by discharging a condenser through highly rarefied gases and vapors, first produced ions of this type, such as the fivefold ionized sulfur ion, the so-called "stripped atoms." They studied these spectra with a vacuum grating spectrograph, Fig. 30. At the present time as many as 23 electrons have been removed from some atoms. These experiments confirmed in a most gratifying way the *spectroscopic displacement law* of Sommerfeld and Kossel, which states that, except for several exceptions explainable by the atomic theory and which will be cleared up later, *the spectrum of any atom is always similar to that of the singly charged ion which follows it in the periodic table (page 16). It is likewise similar to that of the doubly charged positive ion of the element which is two places to the right in the table.* We have shown that this law is theoretically understandable in the special case of the hydrogen-like ions.

The extensive confirmation of this displacement law is of especial significance because it is a proof for the validity of the "construction principle" for the electron shells of the atoms. This principle, originated by Bohr, says that each atom of the periodic table is built up from the atom which precedes it in the table by increasing the nuclear charge by unity and adding one additional electron. If this principle is correct, then, by stripping an outer electron from an atom, we have a positive ion which is similar in its electron arrangement and therefore in its term diagram and spectrum to the preceding element. This is exactly what the experimentally confirmed displacement law states.

3-8. X-ray Spectra, Their Significance in the Atomic Theory, and Their Relation to the Optical Spectra

We showed on page 115 that the spectra of strictly hydrogen-like atoms, those with a nucleus and only one electron, for elements of atomic number above 20, lie in the wavelength region below one angstrom unit, i.e., in the X-ray region. Highly ionized atoms of this type cannot as yet be produced in the laboratory. They are found only in the interior of stars where the atoms, because of the extremely high temperatures, are ionized thermally. The X-rays which we use experimentally must, therefore, come from atoms which retain all or nearly all their electrons. In order to understand the emission of these X-rays which are characteristic of the particular atom, we have to assume several facts about the structure of the electron shells of atoms of higher atomic number. These facts actually have been deduced from the study of X-ray spectra.

a. Electron Shell Structure and X-ray Spectra

As we shall show in detail on page 168 the electrons of an atom of higher atomic number are arranged around the nucleus in individual shells. These shells are called the K, L, M, N, O, and P shell. Only 2 electrons can occupy the K shell; 8 the L shell; 18 the M shell; and in the N shell there is room for 32 electrons, as may be seen from Table 10, page 178. Since the outermost electron shells correspond to the outermost Bohr orbits, and therefore to the highest energy states, we can plot the complete energy level diagram, as in Fig. 65, for the Cu atom with atomic number 29. The K, L, and M shells with principal quantum numbers $n = 1$, 2, and 3, respectively, are filled with electrons, while in the 4-quantum shell there is only one electron, the emitting electron of the copper atom. The optical spectra which can be excited in an electric arc between copper electrodes arise from the excitation of this electron from the 4-quantum state into the normally unoccupied energy states and the subsequent jumping back of the electron into the $n = 4$ state. These higher energy states, which are partially or completely unoccupied, are called optical levels. It can easily be seen that when electrons jump between these outer levels of the copper atom, optical spectra must be emitted or absorbed. Consider one of the outermost electrons of the atom. This is bound to the nucleus by a Coulomb-like electrostatic force, but the high nuclear charge of 29 positive elementary charges is shielded by the inner, occupied electron shells. The outermost electron, available for the optical transitions similar to that of the H atom, moves in a field of a nucleus whose effective charge is of the order of unity. Thus its ionization and excitation energies must be of the same order of magnitude as those of the H atom.

On the other hand, we expect the emission of X-rays as a consequence of transitions between the L and K shells, because here the jumping electron moves in the strong field of either the full nuclear charge, or that charge diminished by the screening of only one unit. Electron transitions between these levels next to the nucleus which give rise to the emission of X-rays are not possible under normal conditions, because all these levels are already occupied by electrons. Emission of X-rays is therefore possible only if there is an unoccupied place in one of the inner electron shells, in other words, if one of the two K electrons is removed from the K shell by ionization. If we neglect the disturbance of the surrounding shells and the screening of the nuclear charge by the other K electron, the energy necessary to ionize one of the K electrons is equal to Z^2R. For the innermost copper electron, since Z equals 29, this energy is about 11,400 ev. It is thus possible to ionize a copper K electron by impact of a cathode ray electron which is accelerated in the X-ray tube by more than 12 kv.

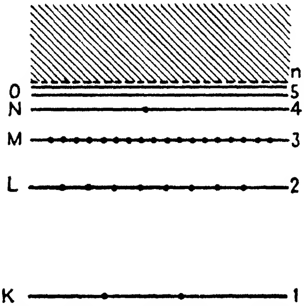


FIG. 65. Simplified energy-state diagram of the Cu atom, indicating also the inner electron states normally filled with electrons which are generally omitted in diagrams referring to optical spectra.

b. The Mechanism of X-ray Line Emission

The mechanism of the emission of X-ray lines, for instance from a copper anticathode, was first conceived by Kossel as follows: Electrons are emitted by an incandescent cathode and accelerated by more than 12,000 volts between the cathode and anticathode of an X-ray tube. The electrons impinging on the copper atoms of the anitcathode penetrate the outer electron shells without collision, and ionize the atom by kicking out one of the two innermost K-shell electrons. To be exact, it is not necessary to ionize completely the K electron in question. It is sufficient to lift it to one of the unfilled optical levels. Since the separation of the unfilled levels from the ionization level is negligible compared with the distance between the L and K levels (greatly distorted in Fig. 65) the entire process is usually called impact ionization.

The hole that is left in the K shell is next refilled, as indicated in Fig. 66, by a transition of an electron from the L, M, or N shell accompanied by the emission of the energy difference in the form of an X-ray line. All X-ray lines whose common end state is the first quantum state are called the K series. Its lines are designated in decreasing order of their wavelength, as K_α, K_β, and K_γ. When the K_α line is emitted, there is an electron hole in the L shell which in turn is filled by an electron from the M or N shell. This transition is coupled with the emission of "softer" X-ray lines, called L_α, L_γ. When the hole in the M shell is filled by transition of an electron from the N shell (which is the highest occupied shell of copper), the very soft M_α line is emitted.

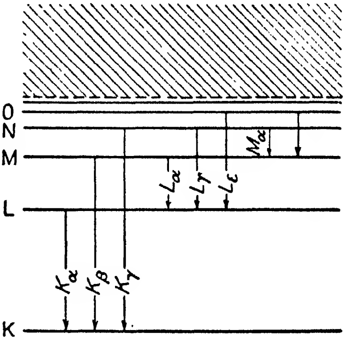


FIG. 66. Energy-level diagram of a many-electron atom with transitions corresponding to the emission of X-ray lines.

Actually the situation is much more complicated since each X-ray line of Fig. 66 consists of a number of closely spaced lines of different intensity. We shall discuss this fine structure of X-ray lines soon.

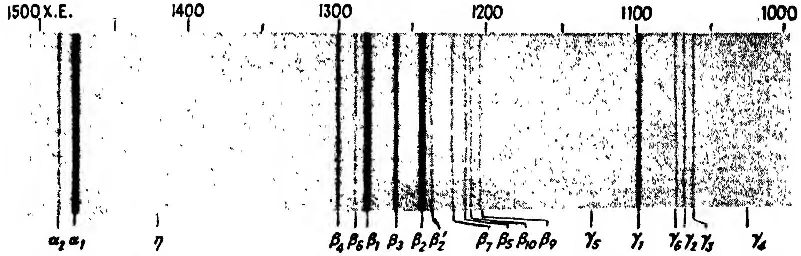


FIG. 67. Example of an X-ray line spectrum: the complete L spectrum of the tungsten atom. (*After Siegbahn.*)

Figure 67 presents a photograph of the L spectrum of tungsten showing all lines. The wavelength unit used here is the XU = 10^{-3} A = 10^{-11} cm. It is evident that the L spectrum can also be emitted without being accompanied by the K spectrum. This happens whenever the cathode-ray electrons do not have sufficient energy to ionize the K shell and thus only L electrons can be ionized by impact.

The X-ray line spectra thus correspond completely to the optical spectra. However, their emission is possible only after the ionization of an inner electron shell makes "room" in a normally filled inner shell for an electron transition. The excitation energy of an X-ray spectrum thus is practically equal to the energy required to throw out an electron from its final state (the ionization energy of the atom from the energy state in question). Thus the wavelengths of the X-ray lines, just as those in the optical spectrum, are determined by the separation of the energy levels (in this case of the innermost occupied ones) and so they are characteristic for any particular atom. In contrast to the continuous X-ray radiation, the X-ray lines are, for this reason, called the *characteristic X-ray radiation.* The greater the atomic number, i.e., the nuclear charge, of a radiating atom, the shorter are the wavelengths of its characteristic radiation, as may be seen from Figs. 6 and 7. We have already used this law, which was discovered by Moseley, in determining the atomic numbers of the elements in Chap. 2. We can understand more clearly the relation between the X-ray lines and the optical lines if we consider the element with the smallest nuclear charge and realize that the K_α line of the hydrogen atom is identical with the longest line Ly_α of the Lyman series of hydrogen (page 101).

The X-ray bremsstrahlung (retardation radiation) which has been discussed briefly on page 111, in striking contrast to the characteristic X-ray radiation, is independent of the material of the anticathode. It is emitted when the cathode-ray electrons are stopped in the nuclear field of the anticathode atoms. The continuum extends, as has already been explained on page 110, from very long waves down to a sharp shortwave limit whose energy $hc\bar{\nu}_G$ is equal to the maximum energy of the retarded cathode-ray electrons. The absolute intensity of the continuum depends definitely on the nature of the retarding atom, in particular, on the atomic number, which determines the strength of the retarding nuclear field. Because of the existence of the X-ray retardation continuum, every X-ray-like spectrum is always superimposed on a continuous background as is nicely shown in Fig. 35.

c. X-ray Absorption Spectra

The processes involved in the emission of characteristic X-ray spectra become especially clear when we consider the question of X-ray absorption spectra. In optical spectra all lines of the resonance series which terminate on the ground state—they correspond to the K series in the X-ray region—can also be absorbed. The absorbing electron then makes a transition from the ground state to a higher energy state. In contrast to this, the absorption of the X-ray lines K_α and K_β is not possible, according to Kossel, because the L and M shells are full and the K electron therefore cannot jump into them by absorbing either line. By absorbing radiation, the K electron can only go to an unfilled level or into the continuous energy region above the ionization limit. If the separation of the optical levels is regarded as negligible compared to the separation of the X-ray levels, then an electron of the K shell can absorb only the K-series continuum, and an electron of the L shell correspondingly only the L-series continuum. *Thus there are X-ray emission lines. but no absorption lines. The X-ray absorption spectrum consists exclusively of the continua beyond the X-ray series, with some structure near the longwave limit of the continuous spectra* which we shall discuss later. These continua are usually called *absorption edges* and they extend from the ionization limit of a particular level with decreasing intensity toward shorter waves.

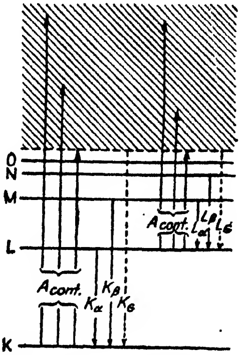


FIG. 68. Energy-level diagram indicating the relation between X-ray emission and absorption spectra.

An absorption process in the continuum of the K series produces again an electron hole in the K shell which gives rise to the emission of the K series. The relation between emission and absorption of characteristic X-rays is illustrated by the transition diagram, Fig. 68. The X-ray lines emitted as a result of the absorption of the X-ray continuum lie on the longwave side of the associated absorption edge. In the case of silver, for instance, the wavelength of the K absorption edge is 0.482 A, whereas the shortwave emission line, as measured, is 0.485 A. Figure 69 is a graph of the absorption coefficient versus wavelength in the vicinity of the K edge; also plotted are the positions of the K_α and K_β emission lines. The absorption coefficient increases suddenly near the ionization limit and gradually decreases toward shorter waves. That the absorption coefficient is not zero on the longwave side of the absorption edge is due to the shortwave extension of the L absorption edge.

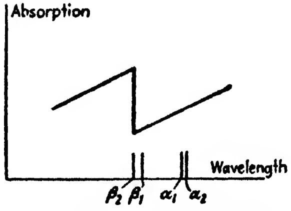


FIG. 69. Schematic representation of the relation between X-ray absorption edge and X-ray emission lines.

Figure 70 is an example of the X-ray absorption edge photographed by Wagner, who simultaneously with M. de Broglie observed and explained these X-ray absorption edges for the first time. If a continuous X-ray spectrum (retardation continuum) is photographed with a silver bromide plate, there is a discontinuity in the photographic density at the location of the absorption edges of Ag and Br. Increasing absorption is indicated by increased blackening. The Ag edge can be easily recognized on the records *a*, *c*, *e*, and *g* in Fig. 70. The increased blackening decreases thus with the absorption from the edge toward shorter waves (left). At *b*, *d*, and *f* an absorbing sheet of Cd, Ag, or Pd was inserted. These sheets absorbed the continuum lying on the shortwave side of their metallic absorption edges. The plate appears less dense at these places. The silver edge coincides, as is to be expected, with the Ag edge of the AgBr emulsion, while the Cd edge is visible on the longwave side and the Pd edge on the shortwave side of the Ag edge.

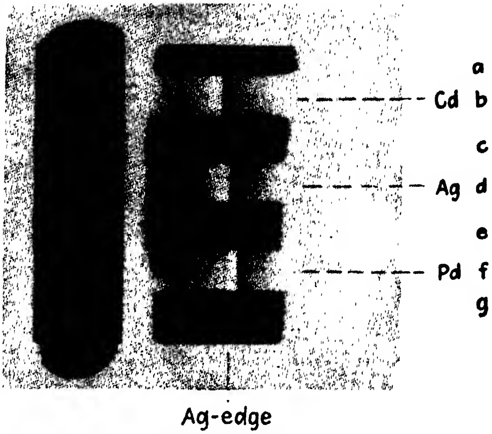


FIG. 70. Photograph of X-ray absorption edges. Detailed discussion in text. (*After Wagner.*)

d. The Fine Structure of X-ray Spectra and Absorption Edges

After this survey of the most essential phenomena of X-ray spectra we conclude with a brief discussion of some more detailed phenomena: namely the structure of the absorption edges and that of the transitions interpreted as single lines in Fig. 68.

We mentioned above that by high dispersion photography complicated structures of the longwave edges of the absorption continua were discovered. According to Kossel, this fine structure depends on the fact that in addition to the true ionization of an inner electron by absorption in the continuum, there is also an excitation possible, that is, a transition of the electron to an unoccupied optical level by absorption of a discrete line. Because the optical term differences are so small—only a few volts, compared to the ionization energy of the inner electrons of many thousand volts—these absorption lines can appear only as a fine structure at the longwave limit of the absorption edge. The existence of these lines has been proved by using argon and nitrogen gas as absorbers of soft X-rays. The case of absorption by solid matter introduces complications since the optical levels here are disturbed and broadened due to the close packing of the atoms, as will be discussed on pages 188 and 438. Moreover, since transitions can occur between optical levels of neighboring atoms, the picture becomes more complicated. In addition to this fine structure, X-ray absorption by solid bodies reveals a secondary structure which in some cases extends over more than a hundred electron volts. According to Kronig, this phenomenon is due to the fact that in a solid body there are no completely free electrons. As will be discussed, page 438, there exist for electrons a number of allowed and forbidden regions of kinetic energy so that the ionization continuum (series continuum) near the ionization limit consists of several individual energy bands (Fig. 203).

We mention in conclusion that the exact position of the longwave edges of the X-ray absorption continua depends upon the binding state of the absorbing atom. For example, the position of the edges of the chlorine absorption in Cl_2, HCl, NaCl, $NaClO_3$, and $NaClO_4$ have various wavelengths because of the different valence of the chlorine atom in these compounds. Since the X-ray spectrum originates in the innermost electron shells, its characteristics are determined essentially by the atomic number, which equals the number of nuclear charges of the atom. However, there is evidence in the X-ray spectra of the influence of outer electron shells of the atom which determine the chemical behavior of the atom.

In our discussion of the emissive transitions, page 119, we have simplified matters to the extent that by the transition of an electron from the L to the K shell only one line, K_α, was emitted. Actually, for each transition there are a large number of lines, as shown in Fig. 67 for the L spectrum of tungsten. As we shall see in considering the optical spectra of the alkali atoms in the next section, only the first quantum state of the atom corresponding to its K shell is single, whereas the L shell actually consists of 3 levels, the M shell of 5 levels, and the N shell of 7 levels. These sublevels coincide for the H atom, but they split up for heavier atoms because of the relativistic motion of the inner electrons in the stronger nuclear fields. This complete diagram of the X-ray terms including the possible transitions resulting in emission lines is presented in Fig. 71. We shall discuss the quantum numbers written on the right side of this figure later. In addition to the prominent lines a number of very weak lines can be seen in the X-ray spectra of the lighter elements. These are called satellites, and their significance is not yet entirely clear. They may be due to the simultaneous excitation of two electrons and the existence of electron shells which are not completely occupied. They are occasionally called

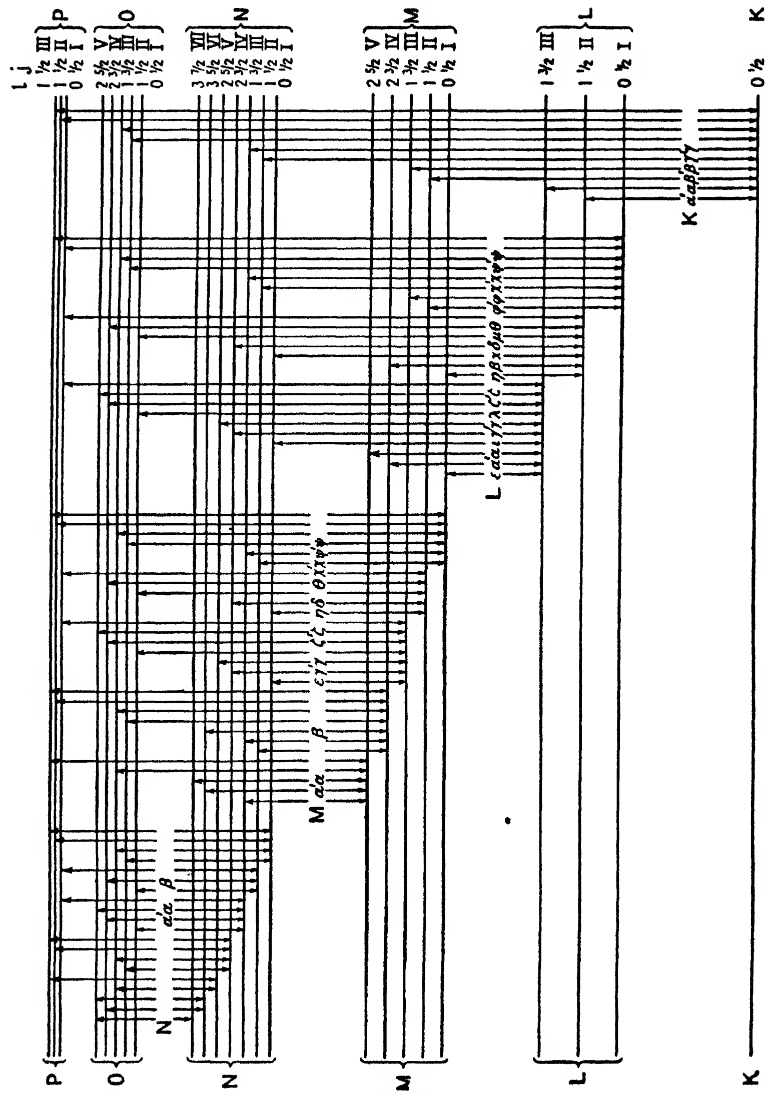


Fig. 71. Complete X-ray-level diagram with transitions and quantum numbers. (*After Sommerfeld.*)

spark lines because it was first assumed that, as in the case of the optical spark spectra (page 113), they were due to electron transitions in an atom which had already lost one of its internal electrons by ionization.

3-9. The Spectra of the Alkali Atoms and Their Significance. S-, P-, D-, F-term Series

Except for the X-ray spectra, we have so far discussed exclusively the spectra of single-electron systems. We now take up the next more complicated group of atoms, the alkalies, Li, Na, K, Rb, and Cs. It is known from their chemical and other behavior that each of these atoms has one or more closed shells and one external electron which can be excited and also ionized fairly easily. In this respect, they are similar to the H atom. The outermost electron, which is called the emitting electron, or, because of its importance for chemical binding, the valence electron, does not move in the simple Coulomb field of the nucleus. The Z-fold positive charge of the nucleus is compensated for by the closed inner shells of electrons, so that the effective charge is of the order of unity. The nucleus with its closed inner electron shells is called the atom core, so one speaks of the core and the emitting electron. *Electrostatically it is not possible to substitute a positive effective charge Z_{eff} for the core. There must be added to its equivalent point charge a perturbation potential which depends on the orbit of the outer electron with respect to the core. This perturbation potential affects the energy states, and thus the spectra, of the alkali atoms in a very characteristic manner.*

We begin again with empirical data. Kayser and Runge had already found that the spectra of each alkali atom could be resolved into three series. Later a fourth series was found in the long-wavelength region. These series are named principal series, sharp series, diffuse series, and Bergmann series. In absorption, at moderate temperatures, only the principal series can be observed. According to page 102, the principal series thus terminates, as a resonance series, on the ground state of the atom. The wave numbers $\bar{\nu}$ of the four series can be represented by Rydberg formulas of the following form:

$$\left.\begin{array}{llll}
\text{Principal series:} & \bar{\nu} = \dfrac{R}{(1+s)^2} - \dfrac{R}{(n+p)^2} & n = 2, 3, 4, \ldots \\[2ex]
\text{Sharp series:} & \bar{\nu} = \dfrac{R}{(2+p)^2} - \dfrac{R}{(n+s)^2} & n = 2, 3, 4, \ldots \\[2ex]
\text{Diffuse series:} & \bar{\nu} = \dfrac{R}{(2+p)^2} - \dfrac{R}{(n+d)^2} & n = 3, 4, 5, \ldots \\[2ex]
\text{Bergmann series:} & \bar{\nu} = \dfrac{R}{(3+d)^2} - \dfrac{R}{(n+f)^2} & n = 4, 5, 6, \ldots
\end{array}\right\} \quad (3\text{-}62)$$

s, p, d, and f are certain constants which are characteristic of the alkali atoms, and R, again, is the Rydberg constant. If we compare formula (3-62) with the Rydberg formula for the hydrogen series, page 101, we see that *for hydrogen the values of s, p, d, and f are zero, and so we conclude that they are due to the disturbance of the emitting electron by the core.* This perturbation is, of course, missing in the case of the H atom. If the value zero is substituted for *s*, *p*, *d*, and *f*, the principal

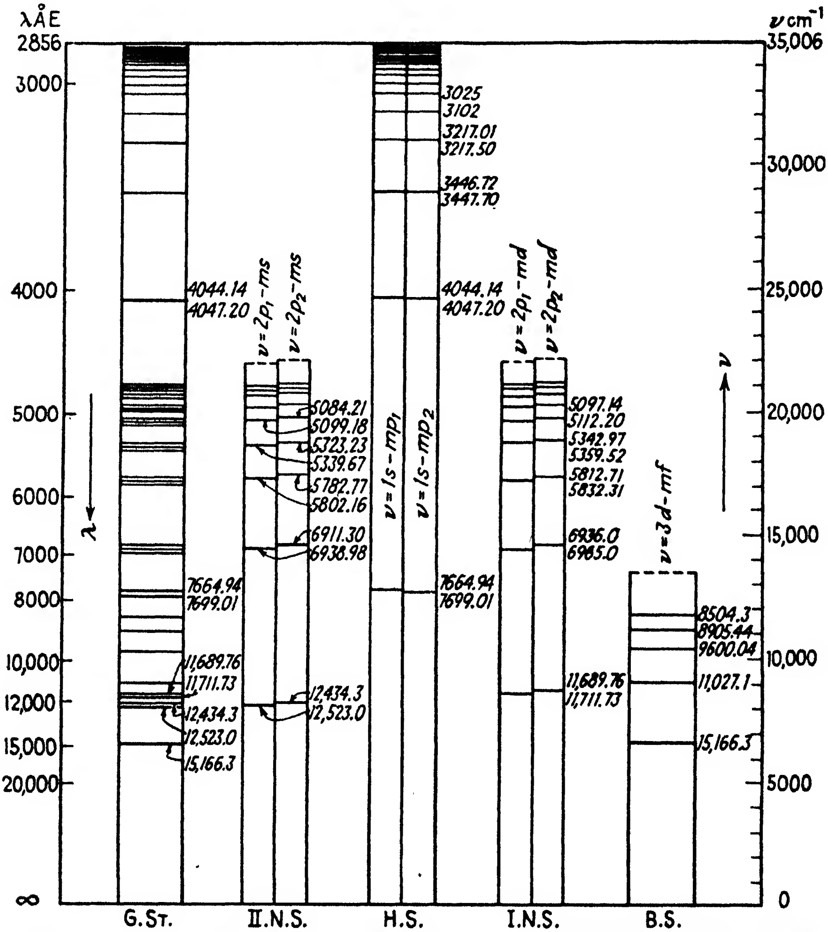


FIG. 72. Schematic representation of the spectrum of the K atom (at the left) consisting of the four series represented at the right. (*After Grotrian.*)

series becomes identical with the Lyman series of hydrogen (3-33), the sharp and the diffuse series then coincide with the Balmer series, and the Bergmann series with the Paschen series. The different series of potassium are shown in Fig. 72 together with its total spectrum resulting from the overlapping of the different series and therefore appearing rather complicated. Figure 73 is a photograph of a part of the emission spectrum of potassium. The last lines of the sharp and diffuse series, which converge to their common limit $2P$, can be seen on the left side of the photograph together with their common series limit continuum. On the right, i.e., toward the shorter waves, the principal series, which converges toward the limit $1S$, can be seen.

From the series formula (3-62) we see that the alkali atoms do not have one single term series as we know it from the H atom, but that there are four series resulting from the disturbance of the emitting electron by the core. The series corresponding to the constants s, p, d, and f are designated by their capital letters. They are represented graphically next to one another in Fig. 74. If the principal quantum number n is added in front of the corresponding symbol of the term series, the Rydberg formulas (3-62) can be written in an abbreviated form:

$$\left.\begin{array}{lll} \text{Principal series:} & \bar{\nu} = 1S - nP, & n = 2, 3, 4, \ldots \\ \text{Sharp series:} & \bar{\nu} = 2P - nS, & n = 2, 3, 4, \ldots \\ \text{Diffuse series:} & \bar{\nu} = 2P - nD, & n = 3, 4, 5, \ldots \\ \text{Bergmann series:} & \bar{\nu} = 3D - nF, & n = 4, 5, 6, \ldots \end{array}\right\} \quad (3\text{-}63)$$

The S state of lowest energy has the principal quantum number $n = 1$; the lowest existing P state, $n = 2$; the lowest D state, $n = 3$; and the lowest F state, $n = 4$. By computing the separation of the energy states from the term analysis of the spectrum according to page 83, an energy-level diagram as shown in Fig. 74 for sodium is obtained. The symbols S, P, D, and F have merely historical significance. P comes from principal series, S and D designate the sharp and diffuse series because of the difference of definition of their lines (which can be explained by the present atomic theory), and F comes from fundamental. Originally the Bergmann series was called the "fundamental series" because a special significance had erroneously been ascribed to it.

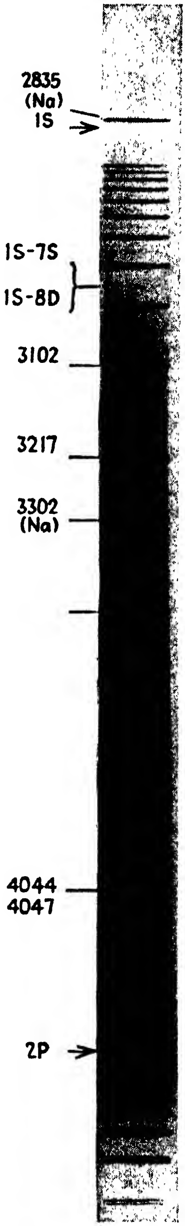


FIG. 73. The spectrum of the potassium atom, as emitted by a potassium-vapor discharge, with the limits of the principal series ($1S$) and the subsidiary series ($2P$). Between the lines of the principal series the intensive series limit continuum of the subsidiary series and two "forbidden" lines ($S \rightarrow S$ and $S \rightarrow D$) can be seen. (*After Krefft.*)

In the case of the hydrogen atom, we have a relation between the principal quantum number n and the term value measured in wave numbers $\bar{\nu}$ per cm (page 98),

$$\bar{\nu} = \frac{R}{n^2} \qquad n = \sqrt{\frac{R}{\bar{\nu}}} \tag{3-64}$$

For an arbitrary atom, which is not similar to the hydrogen atom, we can write a similar expression to define an effective quantum number

$$n_{\text{eff}} = \sqrt{\frac{R}{\bar{\nu}}} \tag{3-65}$$

Using this expression we can compute the n_{eff} values for the various energy states of sodium as shown on the left in Fig. 74. These quantum numbers are no longer integral. We consider the deviation of the effective quantum number from an integral number as a measure of the deviation of that atomic state from hydrogen-like behavior. From Fig. 74 it follows that the higher P, D, and F terms of the alkalies are very similar to hydrogen, while the S terms and the lower P terms are quite dissimilar. This can be understood as follows: The energy states with equal principal quantum number n coincide in the case of the hydrogen atom. In the case of the alkali atoms, the various states, S, P, D, F, are separated as a result of the electrostatic influence of the atomic core on the outer electron. Therefore, it is easily understood that for a highly excited emitting electron the disturbance due to the atomic core is usually negligible. The S states are an exception, because, in the Bohr theory, they correspond to elliptic orbits of great eccentricity which penetrate into the core and are thus considerably disturbed by its field.

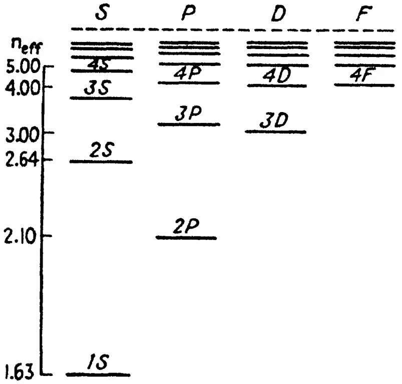


FIG. 74. Energy-level diagram of the sodium atom with the different term sequences and the effective principal quantum numbers of the three lowest terms.

By computing the disturbing potential necessary to separate the S, P, and D terms, it is found that there is a term changing with $1/r^4$ which must be added to the Coulomb term e^2/r. This term can be easily explained. The atom core is polarized by the emitting electron, i.e., it attracts the nucleus and repels the electron shells. The result is an electric dipole with a dipole moment

$$M = \frac{\alpha e}{r^2} \tag{3-66}$$

where α is the polarizability of the atomic core. From the theory of electricity, the potential of a dipole with a moment M is

$$U = \frac{M}{r^2} = \frac{\alpha e}{r^4} \tag{3-67}$$

The disturbing potential, which varies as $1/r^4$ and which is required to explain the alkali spectra, thus is a result of the polarization of the core in the field of the emitting electron. The polarizability of the alkali ions

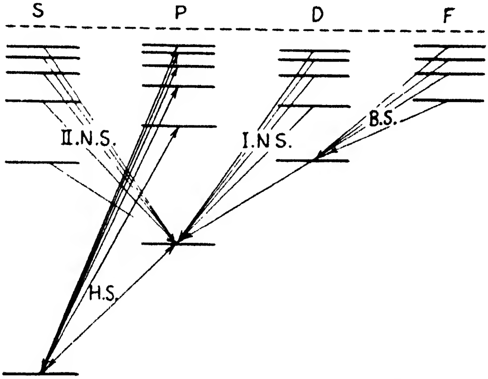


FIG. 75. Term diagram of the potassium atom with the transitions corresponding to the different spectral series. HS = principal series; I NS and II NS = diffuse and sharp subsidiary series, respectively; BS = fundamental series.

computed from their spectra is in good quantitative agreement with the values for the polarizability found by other methods.

By plotting in an energy-level diagram the transitions corresponding to actually observed spectral lines of an alkali atom, using the Rydberg formulas (3-62), we arrive at Fig. 75. We notice that the result is in contrast to the H atom and the original Ritz combination principle. In the alkali spectra, transitions are no longer possible between *all* energy states. Moreover, there is an empirical "selection rule," according to which transitions can occur only between neighboring terms. Transitions between states of the same term series ($S \rightarrow S$, $P \rightarrow P$, $D \rightarrow D$) are "forbidden" just as are transitions between nonneighboring series ($S \rightarrow D$, $F \rightarrow P$). However, we want to point out here that these selection rules are not *absolutely* inviolable, and that forbidden transitions can occur if the emitting or absorbing atom is disturbed by a strong electric field. In discharges of high current density in which line emission of atoms occurs in disturbing fields of neighboring electrons and ions, forbidden lines such as $S \rightarrow S$ and $S \rightarrow D$ occur, although with considerably reduced intensities. Figure 73 shows an example.

In the case of the H atom, an electronic state was unambiguously determined by *one* number, the principal quantum number n. This does not hold for the alkali atoms. In addition to the principal quantum number n, it is necessary here to indicate the special term series with which we are dealing. Sommerfeld introduced for this purpose a second integral quantum number, the secondary quantum number, k. We shall use, instead of k, the orbital quantum number l because, according to Chap. 4, it has a clearer physical significance. l is so defined that for an atom with a single external electron, the S term is characterized by $l = 0$, the P term by $l = 1$, the D term by $l = 2$, etc. Each electronic state (in the Bohr model, each electron orbit) thus is determined by the *two* quantum numbers n and l, where

$$l \leqq n - 1 \tag{3-67a}$$

For labeling the electrons, the small letters s, p, d, f are used following the principal quantum number, so that one speaks of a $1s$ electron, $3p$ electron, etc. Then the selection rule for optical transitions mentioned above states that only those transitions are permitted for which the orbital quantum number l changes by ± 1, or

$$\Delta l = \pm 1 \tag{3-68}$$

In the original Bohr-Sommerfeld pictorial theory of the atom, the principal quantum number n, according to (3-18), determined the diameter of the particular electron orbit and, according to (3-22), the total energy of that electron also. The secondary quantum number $k = l + 1$ determined the eccentricity of the orbit: for $l = n - 1$ the orbit was a circle, for $l < n - 1$ the orbit was an increasingly eccentric ellipse with the larger diameter identical with that of the circular orbit. In this model, the first quantum orbit was always a circular one, since according to (3-67a) there exists only a $1s$ state and no $1p$ state. One circular orbit ($2P$) and one elliptical orbit ($2S$) are associated with the principal quantum number 2. In general, with a principal quantum number n there are associated n orbits of different eccentricity. The name "orbital quantum number" for l is still appropriate, for according to quantum mechanics (Chap. 4) the emitting electron of the alkali atom has a mechanical orbital momentum equal to $\sqrt{l(l+1)}\,(h/2\pi)$. In the case of the H atom, Sommerfeld showed that because of the purely central Coulomb field the energies of the circular and elliptical orbits for the same principal quantum number n are almost exactly equal, so that we need only *one* quantum number, n. The energy states of the H atom, labeled as n, consist in this approximation of n coinciding energy states of different orbital quantum numbers. This is called (n − 1)-fold degeneracy. As a result of the disturbance of the emitting electron by the atomic core, the energy values of orbits having the same n no longer coincide in the case of alkali and other atoms having more than one electron. The nth quantum state then splits up into n different energy states nS, nP, nD, nF, etc. Their separation (energy differences) depends on the disturbance of the emitting electron by the core. It thus increases with the number of electrons in the core. The similarity to hydrogen is greatest in the case of lithium, whereas the departure from the hydrogen-like behavior and the spacing of the states having the same n is greatest for cesium, which has the largest number of electrons.

We shall not carry through quantitatively the computation of the electron orbits by the Bohr theory, by which all details of the alkali terms, including the more or less hydrogen-like behavior of the different states, can be satisfactorily accounted for. For in Chap. 4, in which the further development of quantum physics will be discussed, we shall see that the concept of electrons rotating in orbits does not have the significance which was originally attributed to it.

In our discussion of the alkali spectra and their meaning one important point has not been considered so far. The single lines shown in Fig. 75 are actually closely spaced doublets. This is illustrated by the well-known yellow D lines of sodium, which correspond to the transition $2P \rightarrow 1S$. Actually these are two lines D_1 and D_2 separated by 6 A. The atomic model of the core and emitting electron cannot explain this duplication of energy states of the alkali atom and all other atoms with one emitting electron. It is necessary to further extend this model, which we shall do on page 144.

The spectra of other atoms and ions with *one* emitting electron, such as Cu, Ag, and Au, have the same characteristics as the alkali spectra. The same applies, according to the spectroscopic displacement law, to the positive ions of the earth alkalies Mg, Ca, and Sr, etc., which have lost one of their two outermost electrons by ionization, to the doubly positive ions of the elements in the third column of the periodic table, etc. We have already shown on page 113 that the spectra of these ions are shifted toward shorter wavelengths, because the square of their effective nuclear charge (4 or 9 or, in general, Z^2) appears as a factor in their term formulas (3-51).

At first it seems surprising that the spectra of atoms with *two* external electrons such as helium, the earth alkalies of the second group of the periodic table, as well as the metals Hg, Cd, Zn, display a great similarity to the alkali spectra in so far as they, too, have principal, sharp, diffuse and Bergmann series. On the other hand, they exhibit phenomena (two separate term systems) by which the two-electron spectra differ from the alkali spectra. We shall discuss the explanation of this fact on page 147. The existing similarity can be explained only by the reasonable assumption that *in general in atoms with two equivalent external electrons only* one *is excited and acts as an emitting electron, whereas the other electron can be regarded as part of the core.*

3-10. The Spectra of Many-electron Atoms and Their Term Symbols. Multiple Term Systems and Manifold Excitation

We have already briefly discussed the spectra of atoms with two outermost electrons. We shall now discuss, in general, atoms with several external electrons. Empirically (for example in the Fe spectrum, Fig. 76) two phenomena are evident: an increasing complexity of spectra concomitant with the gradual disappearance of the Rydberg series, and the appearance of several term systems which do not combine mutually.

The empirical term analysis of the helium spectrum, originating from the simplest two-electron atom, proved the existence of a double term system (Fig. 77), with the exception that in the second system on the right the ground term $1S$ is missing. Evidently the two term systems have nothing to do with each other. Lines which correspond to a transition from a term in one system to a term in the other, so-called intercombination lines, do not appear. We shall discuss the theoretical basis for the nonappearance of these intercombination lines on page 147. At first it was believed that there existed two different kinds of helium atoms belonging to the two term systems, and these were called orthohelium and parhelium. That this is not the correct explanation we shall see in discussing the two term systems in the second section below. Investigation of these spectra with higher dispersion and resolution showed, furthermore, that the energy states of parhelium are simple, whereas those of orthohelium consist, with exception of the S states, of three states lying very close together. This different multiplicity of the terms is denoted by writing a superscript to the left of the corresponding symbol (^1S, ^3P). The former parhelium and orthohelium today are called the singlet and triplet system of the He atom, respectively. *This separation of the spectrum into singlet and triplet series* is also found in the cases of the earth alkalies, the metals mercury, cadmium, and zinc, as well as in the spectra of all ions with two external electrons. It is *a phenomenon typical for all two-electron systems.* We might mention here that this effect also occurs with two-electron molecules such as the H_2 molecule, as we shall see in Chap. 6.

One-electron atoms have a doublet-term system, as we showed on page 131 in discussing the example of the sodium D lines. The two-

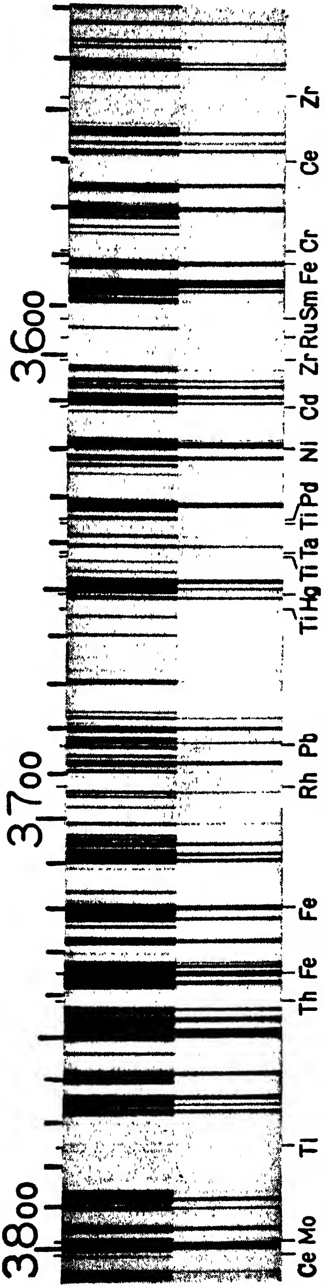


FIG. 76. Portion of the Fe spectrum as an example of a multiline spectrum. Upper part very strongly exposed. Impurity lines indicated below. (*After Gössler.*)

electron atoms have singlet and triplet systems which do not intercombine. For atoms with three outermost electrons, which are found in the third column of the periodic table, we again find, empirically, two term systems which do not intercombine. They are the doublet system and the quartet system, whose terms are twofold or fourfold.

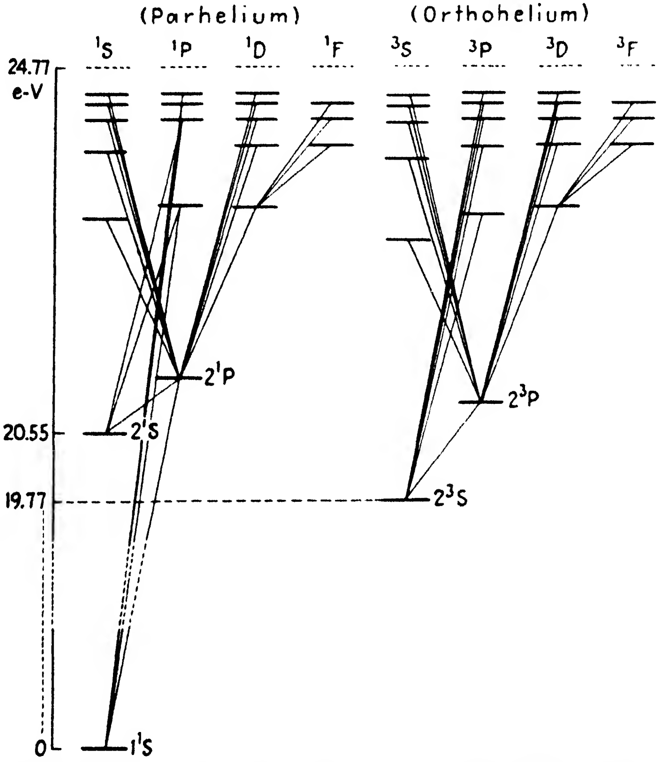


FIG. 77. Term diagram of the helium atom with two entirely separate term systems (singlet and triplet term system).

A four-electron atom has three term systems which do not intercombine, a singlet, a triplet, and a quintet system, while the five-electron atom has a doublet, quartet, and a sextet system. Moreover, each atom has only one ground state ($n = 1$) (for helium 1^1S), which in general belongs to the lowest multiplicity.

The multiplicity of terms and the occurrence of several term systems which do not intercombine, for the many-electron atoms, evidently, is related very closely to the number of outermost electrons. A so-called multiplicity rule has been established from experience: *in going from element to element in the periodic table, even and odd multiplicities* *alternate with each other in such a manner that an odd multiplicity is associated with an atom having an even number of electrons and conversely.*

It has also been empirically established that in the spectra of atoms with the same number of external electrons complexity increases with increasing atomic weight; for example, the spectrum of the light two-electron atom helium can be easily resolved into individual series, in contrast with the spectrum of the heavy two-electron atom mercury, Fig. 38. This difference is caused by the fact that the term separation of the multiplet terms in the light atoms is so small that it appears in the lines only as a fine structure which is difficult to resolve. On the other hand, the term differences in the heavy atoms with a many-electron core, such as Hg, are so large that transitions between these terms lead to widely separated spectral lines, which make the series character of the spectrum extremely complex and unrecognizable. Thus the complexity of the iron spectrum, Fig. 76, is due to the mutual superposition of numerous multiplets. Figure 78 shows another example, a sextet of vanadium, which gives an idea of the complexity of these many-electron spectra. We can now appreciate the considerable difficulty of analyzing these spectra. In this connection a grammatical ambiguity should be pointed out. The expression "multiplicity" (singlet, doublet, or in general multiplet) is used in two senses: for a group of related terms (energy states), as well as for the group of related lines which result from the transitions between such term multiplets. In this second sense, the two D lines of sodium are called a doublet. If one means a related group of terms or energy states, the prefix "term" should always be used.

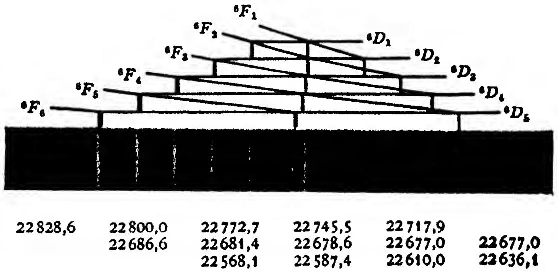


FIG. 78. Sextet in the spectrum of the vanadium atom as an example of a multiplet. (*After Frerichs.*)

The significance of the multiplicity of terms and the accompanying phenomenon of the occurrence of term systems which do not intercombine had for a long time been a difficult theoretical problem. The quantum-mechanical theory of exchange degeneracy (see page 228) has led to the realization that the occurrence of the nonintercombining term systems is due to the identity (i.e., interchangeability) of the electrons of the many-electron atoms. The complete explanation of the systems of different multiplicity was based on the discovery in 1925 by Goudsmit and Uhlenbeck of the electron spin, which we shall discuss in the next section.

However, the complexity of the spectra of atoms with several outermost electrons does not depend exclusively on the multiplicity described above. More important is the fact that in these cases there is no longer a single emitting electron which can be excited and which alone accounts for the energy states of the atom, but there are several. The earth-alkali atom Ca with its two outermost electrons provided the first experimental proof of double excitation. In analyzing the spectrum, a particular term series, the P' series, was found, which does not fit into the normal term system and which behaves peculiarly in many respects. These P' terms converge, as shown in Fig. 79, toward a term limit which lies above the ionization limit. In order to separate the electron responsible for the P' terms, an amount of energy ΔE_i in excess of the ionization energy of the normal atom is required. However, since the ionization must be independent of the way in which it is produced, this amount of energy can mean nothing else than that in the ionization of the P' term series the Ca^+ ion is not in the ground state but in an excited state. Actually the energy difference ΔE_i is equal to the first excitation potential of the Ca^+ ion. Consequently, the P' terms of the earth-alkali atoms correspond to energy states in which one electron is in the first excited state while the second can be excited to any energy state. In the transition from a P' term to an unprimed term both electrons must change their energy states (in the old conception they must change their orbits). That such a double jump is possible is evidence for the relatively strong mutual coupling of the two electrons.

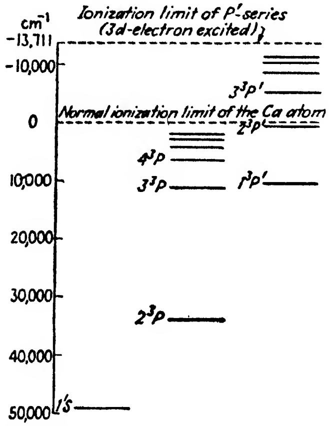


FIG. 79. Term diagram of the Ca atom with P' term sequence of the doubly excited Ca atom converging toward a limit which lies above the normal ionization limit by an amount equal to the excitation energy of the second electron.

In the case of atoms with several outermost electrons, this coupling has the effect that the Rydberg series, so prominent in the spectra of hydrogen and the alkalies, tend to vanish. When a certain amount of excitation *energy is supplied, it seems more probable that several electrons are excited than that one is excited to a very high energy state. Manifold excitation and multiplicity of terms are thus responsible for the confusing abundance of lines in the spectra of many-electron atoms.*

After this survey of the spectroscopic phenomena in the many-electron spectra, we must consider in greater detail the coupling of the different electrons and the character of the terms resulting from this coupling.

We have characterized each electron by assigning to it a principal quantum number, and the symbol s, p, d, or f for its orbital quantum number. An energy state of an atom with one electron is described sufficiently by the symbols of the emitting electron, which determines the entire state of the atom. The quantum numbers of the atomic terms for the many-electron atoms can be determined from the spectra, and in some cases from the influence of a magnetic field. The different possible term characters of an atomic energy state are symbolized by the capital letters S, P, D, or F. *These term characters are determined by the arrangement and behavior of all the outermost electrons of the many-electron atom.* In the Bohr theory the state of the atom depends on the resulting orbital momentum $\mathbf{L}$ of all electrons. The S, P, D, F terms of an atom correspond to the resulting orbital momentum 0, 1, 2, 3, measured in units of $h/2\pi$. This will become clear on page 143. Because the alkali atoms are one-electron atoms, the term quantum number L in this case can be identified with the orbital momentum $\mathbf{l}$ of the one emitting electron, measured in units of $h/2\pi$. Therefore, it makes no difference in the case of the one-electron atom whether the whole atomic state is characterized by the capital letters S, P, D, F, or its single electron is denoted by the small letters s, p, d, f. What is the situation with many-electron atoms, especially those with several electrons in the external shell? Evidently the total orbital momentum of the electron shell $\mathbf{L}$ must be the vector sum of the orbital momenta $\mathbf{l}_i$ of the individual electrons. By adding vectorially the $\mathbf{l}_i$, we can disregard the closed electron shells entirely, since their resulting orbital momentum is always zero. This is not only theoretically reasonable (page 174), but follows empirically from the fact that the ground states of all noble-gas atoms (which are characterized by closed electron shells) are ${}^1S$ states and thus have a resultant orbital momentum $\mathbf{L} = 0$. We have, consequently, to consider only the orbital momenta $\mathbf{l}_i$ of the valence electrons in the outer shell. Their vectorial sum, i.e., their resultant $\mathbf{L}$, depends on the mutual orientation of their orbital planes.

The quantum theory requires that not only the individual orbital momenta $\mathbf{l}_i$, but also the resultant orbital momentum $\mathbf{L}$, be quantized, that is, they must be integral multiples of $h/2\pi$. Consequently, not every arbitrary orientation of the individual electron orbits is possible. They must be so arranged that the resulting momentum is integral. Two p electrons, each of which has the orbital momentum $\mathbf{l} = h/2\pi$, can, therefore, produce an S, P, or a D term of the atom corresponding to $L = 0$, 1, or 2. Figure 80 shows how the individual $\mathbf{l}$ vectors add up to give the three possible atomic states. In the same manner, by adding vectorially the individual orbital momenta of all electrons, the possible atomic terms can also be determined for a larger number of valence electrons.

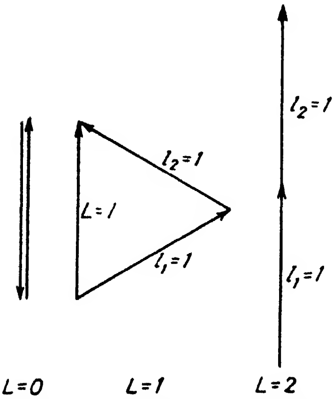


FIG. 80. The three possibilities of vectorial addition of the orbital momenta of two p electrons ($l = 1$) forming the resulting orbital momentum $\mathbf{L}$ of the total electron cloud, corresponding to the S, P, and D state of the atom, respectively.

In order to characterize an atomic state quite precisely, the symbols of all electrons, or at least of all external valence electrons, are written in front of the symbol of the whole atomic state, which in turn has as a superscript on the left the multiplicity of the state. Several equivalent electrons are denoted by writing the corresponding number as a superscript on the right. For example, the symbol for three $2p$ electrons is $2p^3$. The ground state of the helium atom would be written accordingly as $1s^2\,{}^1S$; an excited helium state of the triple system, $1s\,3p\,{}^3P$. In order to abbreviate, the symbol for the individual electrons is frequently omitted, and only the principal quantum number of the highest electron is written before the term symbol. Thus the last-named helium state is written 3^3P.

The individual electrons interact electrostatically as well as through magnetic fields which they set up in rotating about their orbits. As a result of this coupling of the electrons, the whole atom acts as a system of coupled gyroscopes, since each electron rotating in its orbit forms a small gyroscope. Such a system is governed by the law of conservation of the angular momentum, which requires that the resultant angular momentum remain constant in space with respect to magnitude and direction, if no external forces are acting. The individual orbital momenta of the electrons, $\mathbf{l}_i$, must, therefore, precess about $\mathbf{L}$ (Fig. 81). The angular velocity of this precession is the larger, the greater the interaction of the individual electrons. If the interaction of the individual electrons is sufficiently large, the velocity of precession may be of the same order of magnitude as the velocities of the individual electrons in their orbits. In that case, the individual momenta lose their physical significance, and only the total orbital momentum **L** which determines the term character always retains its meaning, that is, it is exactly defined. We shall refine this picture in the next section when we discuss the angular momenta of the electrons around their own axes.

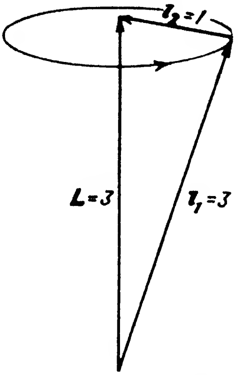


FIG. 81. The precession of the orbital momenta $\mathbf{l}_i$ of the single electrons about the resulting orbital momentum **L** of the atom. Here, one *s* electron and one *f* electron form an *F* state of the atom.

We had established the selection rule for transitions of the one external electron of the alkali atoms to be

$$\Delta l = \pm 1 \qquad (3\text{-}68)$$

This selection rule also remains valid for transitions of each single electron in the many-electron atom. There is, however, a selection rule which is valid for the resulting orbital momentum **L**, which requires

$$\Delta L = 0 \text{ or } \pm 1 \qquad (3\text{-}69)$$

This means that transitions without a change in L are possible. Because of (3-68) these occur only in the case where the transitions of two electrons occur simultaneously so that their orbital momenta change in such a manner that L remains constant. The selection rule $\Delta L = 0$ thus holds only for heavy atoms with strong electron interaction, where several electrons can simultaneously change states.

3-11. Metastable States

Before we discuss the theory of multiplets and the phenomena associated with them, we shall briefly discuss an important consequence of the existence of term systems which do not intercombine: the appearance of metastable states and their importance in atomic physics.

From the term diagram of the helium atom, Fig. 77, it can be seen that the lowest state of the triplet system, the 2^3S state, lies 19.77 ev above the 1^1S ground state of the atom. The transition $2^3S \rightarrow 1^1S$ with the emission of the energy as a spectral line does not occur: intercombinations are forbidden. A He atom in the 2^3S state thus cannot give up its high excitation energy of almost 20 ev by radiation. This energy can be given up only in a collision of the second kind, as discussed on page 104. We have already mentioned that this type of state, whose combination with the ground state is forbidden, is called metastable, in contrast to the normal excited state which, after about 10^{-8} sec, spontaneously returns to the ground state, either directly or in steps. The metastable states cannot be excited from the ground state by absorption of radiation since this, as the reversed process of emission, is also forbidden. However, metastable states can be excited by collisions (electron collisions or collisions of the second kind), since there are no sharp selection rules for collisions.

Under certain circumstances there exist other metastable states. In the case of helium the 2^1S state with an excitation energy of 20.55 ev (Fig. 77) is metastable, since the selection rule $\Delta l = \pm 1$ forbids the transition $2^1S \rightarrow 1^1S$, and a lower P state, to which a transition might occur, does not exist. In the case of the alkalies, according to Fig. 75, the 2^2S state can go over to the lower 2^2P state, and this again to the 1^2S ground state, so that there are no metastable states for the alkalies. *Quite generally, metastable states are excited states of atoms or molecules whose combination, direct or indirect, with the ground state is forbidden.*

Metastable states were first found in electron impact experiments by which, according to Fig. 50, page 91, critical excitation potentials were measured for which no corresponding line in the spectrum could be found. The law forbidding radiation from a metastable state is not an unbreakable one. All selection rules are valid, as was mentioned on page 129, but with certain restrictions. In the current treatments of electrodynamics, only the radiation of electric dipoles is discussed, and this dipole radiation is not possible from metastable states. However, there are other possibilities of radiation, for example, by electric quadrupoles or by magnetic dipoles, which are possible for metastable states. The intensity of this radiation compared to that of normal dipoles is extremely small. Moreover, as we have already mentioned, forbidden transitions become less improbable if the electric field of neighboring electrons and ions is sufficiently strong. A perturbation of the atom which can be called an internal one in contrast to the external disturbance due to environment, occurs, for instance, in heavy atoms such as mercury. Here the disturbance of the two external electrons by the core is so strong that singlet-triplet intercombinations, which are strongly forbidden for the lighter atoms, appear with considerable intensity. The well-known Hg line 2537 A is such an intercombination line. Such deviations from the selection rules and the existence of quadrupole and magnetic dipole radiation make possible the emission of excitation energy which has not yet been lost by collisions. For this reason, the mean life of metastable states is by no means infinitely great, but it is large compared to the mean life of normal excited states of about 10^{-8} sec. Meissner and Dorgelo in 1925 made the first direct measurements of the mean life of excited states. For metastable states, they obtained a value of the order of magnitude of 10^{-2} sec. However, metastable states exist which have a life of the order of 1 sec.

In general, if investigations are not carried out in a very high vacuum, the mean life of metastable atoms in gas discharges is limited by collisions of the second kind (page 104), and not by radiation. Only under the most extreme conditions of gaseous and planetary nebulae of extremely small density and in highly rarefied atmospheres (such as in the upper layers of the earth's atmosphere) is it possible to have undisturbed radiation from metastable atoms. According to Bowen, the mysterious "nebulium lines," which had perplexed spectroscopists for many decades, can be explained by such forbidden transitions occurring in the ions O^+, O^{++}, and N^+. In the meantime, the disputed lines of the aurora borealis were explained by McLennan and Paschen as due to similar forbidden transitions, this time within the neutral O atom.

We have been using the term "metastable atom" for an atom in the metastable state. This expression has come into use because atoms in the metastable state have large amounts of energy which they can give up in collisions of the second kind, and because, in addition to normal atoms and ions, they are important constituents of each gas discharge (plasma). Furthermore, the mean life of the usual excited states of atoms is so small, 10^{-8} sec, that these can exist only in comparatively small concentrations. They are thus negligible compared with metastable atoms. Many phenomena in gas discharges which at first were not understood have been explained as being due to metastable atoms. Carrying energy, they can diffuse from one region of the discharge tube to another without being influenced by an electric field as are positive ions.

Metastable atoms cause ionization of other gases mixed with them, if their excitation energy is larger than the ionization energy of the gas. Metastable He atoms with high energies of about 20 ev, as well as those of the other noble gases, are especially effective when mixed with vapors of metals which have relatively low ionization energies (see Table 2, page 30). For example, this ionizing effect will lower the breakdown potential required to strike a neon discharge, if a small amount of Hg vapor is added to the neon. The Hg atom is ionized in a collision of the second kind with the metastable neon atom, and acts as a new charge carrier.

If the energy of the metastable atoms is not sufficient to ionize the admixed gas, the metastable atoms can diffuse to parts of the apparatus which are not exposed to accelerated electrons or shielded by an electric field, and can then excite the gas or vapor. Of special interest is the possibility of exciting atoms or molecules to a definite energy level which corresponds to the energy of the exciting metastable atom.

There is also a very high probability that metastable atoms excite molecules, causing them to dissociate into their atoms or atom groups (see Chap. 6, page 376). This is due to the fact that molecules have a large number, and therefore a large density, of energy states (including vibrational and rotational states), so that, compared to conditions which exist in atoms, there is a high probability for resonance between metastable atoms and molecules.

A fourth effect of metastable atoms is the release of secondary electrons from metal surfaces (page 474). Such secondary electrons, which may also be released from metal parts of a tube other than the actual electrodes, can be responsible for the many disturbing effects in gas-discharge measurements.

In concluding we mention that metastable atoms play an interesting role in discharges of high current density, such as in high-current arc columns, where, as was discovered by the author, they act as an admixed gas of extremely small ionization potential. They effectively reduce the anode drop of these discharges when they diffuse out of the column into the anode drop region.

It has already been mentioned that metastability is not limited to atoms but is also found in molecules. Metastable nitrogen molecules are especially well known because of their role in the mechanism of the surprisingly long afterglow of certain nitrogen discharges, the so-called active nitrogen.

3-12. Electron Spin and the Theory of Multiplets

We return now to the atomic theory proper and ask what characteristics an atomic model must have in order to explain the actually observed emission and absorption. We explained the spectrum of the H atom using the Bohr theory, page 97, by the stationarity condition (electrostatic attraction equal to the centrifugal force) and the quantization of the momentum p of the electron, which led to the introduction of the principal quantum number n. To explain the splitting of the terms of the alkali atoms into the different term series (S, P, D, F) it was necessary to introduce a second quantum condition, the quantization of the orbital momentum $\mathbf{l}$ of the electron. We now must refine and extend our atomic model to explain the spectra of the many-electron atoms, especially their various term systems and their multiplicity. This became possible when in 1925 Goudsmit and Uhlenbeck, after all other attempts to explain the multiplicity had failed, discovered that the electron itself had an angular momentum (spin) and an associated magnetic moment (page 36). The pictorial concept of the electron, which is open to further refinement, is that in addition to its orbital motion around the nucleus, the electron rotates about its own axis just as the earth rotates simultaneously about the sun and its own axis. The rotation of a charged sphere about its axis is equivalent to a circular current which gives rise to a magnetic field whose direction is that of its axis. Consequently, in addition to the mechanical spin momentum, the electron must also have a magnetic moment associated with the spin.

This magnetic moment is called the Bohr magneton. Its magnitude, as given on page 36, is

$$\mu_B = \frac{eh}{4\pi m_e c} = 9.27 \times 10^{-21} \quad \text{gauss cm}^3 \tag{3-70}$$

This value (with the small correction mentioned on page 36) follows from the magnetic properties of the atoms and the free electron to be discussed on page 152. From the atomic spectra, Goudsmit and Uhlenbeck concluded that the spin must have the value

$$\mathbf{s} = \frac{1}{2}\frac{h}{2\pi} \tag{3-71}$$

Let us consider what this means. The Bohr quantum condition for the principal quantum number is, according to Eq. (3-17),

$$\oint p\,dq = 2\pi r p = 2\pi m r^2 \omega = nh \qquad n = 1, 2, 3, \ldots \tag{3-72}$$

Now if we write this equation

$$pr = mr^2\omega = J\omega = n\frac{h}{2\pi} \tag{3-73}$$

where J is the moment of inertia of the atom, we see that actually it is not the momentum p of the electron, but the angular momentum $J\omega$ of the atom, which is quantized, i.e., it is always a multiple of $h/2\pi$. In the quantum theory this quantity $h/2\pi$ is often designated by the symbol h and is called "h bar." Until now we have always made our computations on the basis of circular orbits. If, however, the orbits are elliptical, then both r and φ vary simultaneously. For the orbital momentum $\mathbf{l}$, in which only φ varies, we can then write an expression which corresponds to the quantum condition (3-17) or (3-72)

$$\oint |\mathbf{l}|\,d\varphi = 2\pi\,|\mathbf{l}| = lh \qquad l = 0, 1, 2, \ldots \tag{3-74}$$

Here

$$|\mathbf{l}| = l\frac{h}{2\pi} \tag{3-75}$$

where l is the orbital quantum number which we had to introduce on page 130 in order to explain the alkali spectra. The orbital momentum of an electron rotating about the nucleus is represented by a vector drawn perpendicular to the plane of the orbit at its center (Fig. 82). Its length is equal to the orbital momentum and its direction is that of a right-handed screw rotating in the same sense as the electron.

According to Goudsmit and Uhlenbeck, the spin momentum of the electron is to be measured in the same units, $h/2\pi$. In order for its value to be in agreement with the spectra, it must, according to (3-71), have the constant value $\frac{1}{2}(h/2\pi)$. In other words, we can say that the spin quantum number is constant, $s = \frac{1}{2}$.

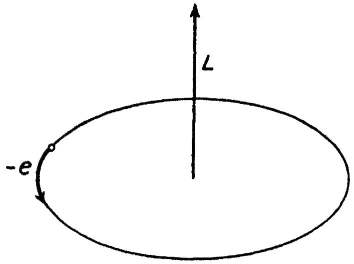


FIG. 82. Orbital momentum and orbital motion of an electron.

In our treatment of the spectra, so far, we have considered only the orbital momentum of the electron, whereas actually *each electron of an atom has two momenta associated with it, the orbital momentum* $\mathbf{l}$ *(having different values for the s, p, d, f electrons according to page 130) and the spin momentum with the constant value* $\mathbf{s} = \frac{1}{2}(h/2\pi)$. *The two momenta are coupled through their corresponding magnetic fields, and combined vectorially constitute the total momentum of the electron.* According to Sommerfeld, all phenomena connected with the multiplicity of the spectra of many-electron atoms can be explained by assuming that the total momentum $\mathbf{j} = \mathbf{l} + \mathbf{s}$ of the atomic electron is quantized as follows:

$$|\mathbf{j}| = j\frac{h}{2\pi} \qquad j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots \tag{3-76}$$

This means that, due to quantum conditions, only those orientations of the orbital momentum $\mathbf{l}$ and the spin $\mathbf{s}$ are possible, for which the corresponding values of the total quantum number j differ by unity. That the j values in (3-76) are half-integral is due to the fact that the spin quantum number s is always $\frac{1}{2}$.

We begin by discussing the simplest case, that of the alkali atoms with only one outermost electron. In the case of the S term of the alkalies the orbital momentum $\mathbf{l}$ is zero, so that there is no magnetic field due to the electron rotating in its orbit with respect to which the spin $\mathbf{s}$ could orientate itself. For that reason S terms are always single. In the case of a P term, the orbital momentum $\mathbf{l} = 1$ couples with the spin $\mathbf{s} = \frac{1}{2}$ of the emitting electron, so that the values of the total momenta differ by integral numbers (always in units $h/2\pi$). It is evident that there are two possibilities of satisfying this condition, namely

$$\left.\begin{aligned} j_1 &= 1 + s = \tfrac{3}{2} \\ j_2 &= 1 - s = \tfrac{1}{2} \end{aligned}\right\} \tag{3-77}$$

Because of the existence of the electron spin $\frac{1}{2}$, the P states of the alkali atoms are split into two states which are characterized by the total momentum quantum numbers of the electron shells, $j = \frac{3}{2}$ and $j = \frac{1}{2}$. In the case of the D terms with $l = 2$ we obviously have the two possibilities

$$\left.\begin{aligned} j_1 &= l + s = 2 + \tfrac{1}{2} = \tfrac{5}{2} \\ j_2 &= l - s = 2 - \tfrac{1}{2} = \tfrac{3}{2} \end{aligned}\right\} \tag{3-78}$$

Except for the S terms, all terms of the alkali atoms, because of the two possible orientations of the spin, are doublet terms. The S terms also can be regarded as doublets, but because $l = 0$, the spin orientation is ineffective, so that the S terms appear to be single.

The influence of the electron spin of constant value $\frac{1}{2}(h/2\pi)$ is the reason for the doublet character of the alkali terms, which is in agreement with the spectroscopic results. The quantum number j of the total momentum by which the two doublet components differ is written as a subscript, at the right of the term symbol. The two doublet components of an alkali atom associated with $n = 3$ and $l = 1$ thus are written $3^2P_{\frac{3}{2}}$ and $3^2P_{\frac{1}{2}}$, and the corresponding ground state $1^2S_{\frac{1}{2}}$.

The energy order of the term components of different j values in general is such that the term with the higher j value also has a higher energy than a term of lower j value (so-called "normal terms"). Sommerfeld also found, however, "inverted terms" like those of cesium, where the $^2F_{\frac{7}{2}}$ term has a lower energy than the $^2F_{\frac{5}{2}}$ term. Because of the increasing interaction between the electrons, the spacing of the doublet states increases with increasing atomic weight; e.g., it is small in the case of lithium and very large for cesium. It is, furthermore, greater for the P terms than for terms with higher l or L values, respectively.

From the spectra we derive for the total momentum quantum number j the selection rule for allowed optical transitions (which follows automatically from quantum mechanics, Chap. 4),

$$\Delta j = 0 \text{ or } \pm 1 \tag{3-79}$$

with the exception that the transition from $j = 0$ to $j = 0$ is forbidden.

From this transition rule and the section of the alkali term diagram,

Fig. 83, which, in contrast to Fig. 75, also indicates the doublet separation of the P term, we see that all $P \rightarrow S$ transitions are double, resulting in doublet lines such as the well-known sodium D lines at 5890 and 5896 A. In the case of transitions between P, D, and F terms, Fig. 84, we should theoretically expect three lines, since for $D \rightarrow P$ the transitions $\frac{5}{2} \rightarrow \frac{3}{2}$, $\frac{3}{2} \rightarrow \frac{3}{2}$, and $\frac{3}{2} \rightarrow \frac{1}{2}$ are allowed. However, a selection rule greatly favors the $\frac{5}{2} \rightarrow \frac{3}{2}$ and $\frac{3}{2} \rightarrow \frac{1}{2}$ transitions compared to the third transition $\frac{3}{2} \rightarrow \frac{3}{2}$ which, because of its low intensity, is called a *satellite*. Only by neglecting the satellite, the three transitions therefore appear as a doublet. The spectra of the one-electron alkali atoms are thus explained.

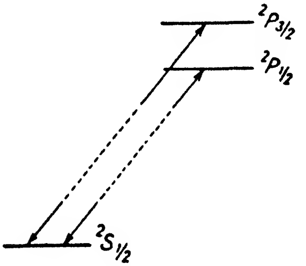


FIG. 83. Schematic term diagram and transitions producing a $^2P \rightarrow {}^2S$ doublet of an alkali atom.

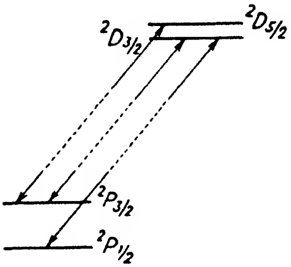


FIG. 84. Schematic term diagram and transitions producing a $^2D \rightarrow {}^2P$ doublet of an alkali atom.

In the case of the many-electron atoms, according to page 137, we have to add vectorially the momenta of all external electrons to obtain a total momentum $\mathbf{J}$ of the atomic shell. This now contains, however, in contrast to page 137, contributions of the spin momenta of the electrons. The computation of the resultant total angular momentum of an atom from the orbital and spin momenta of its external electrons depends on the strength of coupling between the orbital and spin momenta of each electron, on the one hand, and of that between all the angular momenta, or the spins of all electrons among themselves, on the other hand.

We know from the study of atomic spectra that in the majority of all atoms, and without exception in atoms which are not too heavy, the orbital and spin momenta of the outer electrons follow the so-called Russell-Saunders LS coupling. In this case, the interaction between all the orbital momenta $\mathbf{l}_i$ of the electrons, and that between all spins $\mathbf{s}_i$, is large compared with the interaction between the orbital momentum and the spin momentum of the individual electrons. Thus the orbital momenta $\mathbf{l}_i$ of all outermost electrons combine to form the resulting orbital momentum **L** of the atom, which determines the term characters (*S*, *P*, *D*, or *F*). All spin momenta, $\mathbf{s}_i = \frac{1}{2}$, of the individual electrons, on the other hand, combine vectorially to form a resultant spin momentum **S** of the electron shell, which for an even number of electrons is integral and for an odd number of electrons is half-integral. Finally, **L** and **S** combine vectorially to form a quantized total momentum **J** of the atom, about which **L** and **S** precess because of their gyroscopic properties.

In the second coupling case, the so-called *jj* coupling, which occurs predominantly in the case of the excited states of the heaviest atoms, the interaction of the orbital momentum **l** and the spin **s** of each individual electron is large compared with the interactions of the $\mathbf{l}_i$ among each other and of the $\mathbf{s}_i$ among each other. The orbital momentum and the spin of each electron combine to form a total momentum **j** of the electron, and the various $\mathbf{j}_i$, in turn, are to be added vectorially to form the resultant quantized total momentum **J**, about which they precess. The essential point is that in the *jj* coupling the resultant orbital momentum **L** is not defined any more, and there is no possibility of designating the terms by symbols *S*, *P*, *D*, or *F*. In the case of *jj* coupling a term is labeled by its *J* quantum number, which is also defined in the case of the Russell-Saunders *LS* coupling, but there it is of less significance than the *L* value in designating the energy of the term. In the following we shall go into some details only for the more important cases of *LS* coupling. For details of the *jj* coupling and the transition from *LS* coupling to *jj* coupling the reader is referred to special spectroscopic texts.

The selection rule for the total quantum number *J*, which limits the possible optical transitions in a many-electron atom, is identical with that valid for the total momentum quantum number *j* of the single electron,

$$\Delta J = 0 \text{ or } \pm 1 \tag{3-80}$$

again with the restriction that the transition 0 → 0 is forbidden.

We discuss now the results of *LS* coupling with respect to the multiplicity of the various many-electron atoms. In the case of a two-electron atom, such as helium, we have two possibilities for the orientation of the electron spins: parallel orientation which gives $S = 1$, antiparallel, $S = 0$. We shall show that the first case corresponds to the triplet-term system, the second to the singlet-term system of the helium atom. *That intercombinations between two term systems are forbidden, i.e., never occur in the case of light atoms and in the case of heavier atoms only with small intensities (page 140), evidently means that a reversal of the electron spin from* $+\frac{1}{2}$ *to* $-\frac{1}{2}$, *even combined with a change of other electron properities, is very improbable.* This difficulty of a reorientation of the electron spin (change from singlet to triplet terms) accounts also for the different shapes of the electron excitation functions of singlet and triplet levels (Fig. 55) mentioned on page 96. The ground state of the atom is a singlet state. The excitation of a triplet state (spins parallel) from the singlet ground state (spins antiparallel) by electron collision is possible only if the colliding electron replaces one of the two atomic electrons (the one with an oppositely directed spin!). This exchange requires more time than a simple energy transfer and so does not occur if the colliding electron has too much kinetic energy. The excitation function of triplet lines, according to Fig. 55, page 94, decreases beyond the maximum much faster than that of singlet lines.

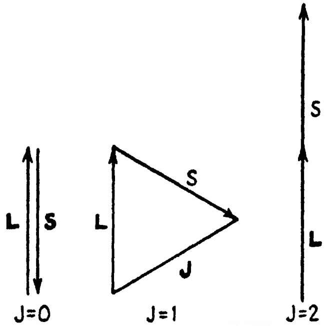


FIG. 85. The three possibilities of vectorial addition of the resulting orbital momentum **L** = 1 and the resulting spin **S** = 1 forming the resulting total momentum **J** of the atom.

Now we consider the terms corresponding to $S = 0$ and $S = 1$. Since, in the case of S terms, both electrons of the atom have an orbital momentum $l = 0$, there is no orbital magnetism (page 154) and therefore no possibility of different orientations even in the case of the total spin $S = 1$; S terms are always single. If one of the two helium electrons has an orbital momentum 1 (p electron), then we have $L = 1$ and a P term. In the case with resulting spin $S = 0$ it is evident that the P terms (since there is no resultant spin) are single. The same holds true for the D and F terms. For resulting spin equal to zero all terms are singlet terms. In the second case, $S = 1$, there are three possible orientations of **L** and **S** whose resultants **J** differ by unity. Figure 85 shows that these lead to the values 0, 1, and 2 for the resultant total momentum **J** of the helium atom. All P states which are single states if the electron spin is neglected are, for the resultant spin $S = 1$, separated into the three states $^3P_0$, $^3P_1$, and $^3P_2$. The resultant spin $S = 1$ thus leads to the triplet system of the He atom. In the same way for the case of the D terms with $L = 2$, combined vectorially with $S = 1$, we have the three terms $^3D_1$, $^3D_2$, and $^3D_3$.

The resultant spin $S = 0$ thus produces a singlet term system with multiplicity 1; the spin $S = \frac{1}{2}$ (alkali atoms), a doublet term system with multiplicity 2; and the spin $S = 1$, a triplet term system with multiplicity 3. From these examples we can see that the expression relating the multiplicity of the terms with the total spin S of the electron shell of an atom can be written

$$\text{Multiplicity} = 2S + 1 \tag{3-81}$$

When there are three outermost electrons we have, as possible values of the total spin of the atom, $S = \frac{1}{2}$ and $S = \frac{3}{2}$, and consequently doublet- and quartet-term systems. Atoms with four outermost electrons may have the S values 0, 1, and 2 and consequently singlet-, triplet-, and quintet-term systems, etc. In the case of the vanadium atom with five outermost electrons, a sextet of which is shown in Fig. 78, the total spin S may have the values $\frac{1}{2}$, $\frac{3}{2}$, and $\frac{5}{2}$, and the spectrum reveals three term systems with doublet, quartet, and sextet terms. From Fig. 86 and Table 7 it can be seen that this highest term multiplicity, six, can be obtained by combining **L** and **S** only for $L \geqq 3$, i.e., F terms. From Fig. 87 it follows that, in agreement with the J selection rule (3-80), the 14 lines marked in Fig. 78 correspond to the combination ${}^6F \longrightarrow {}^6D$. In Fig. 78 the upper terms are represented as oblique lines and the lower terms by horizontal straight lines, as was first done by Meggers; each point of intersection corresponds to a spectral line, as shown in the photograph. Such analysis of multiplets is a very important task of spectroscopists.

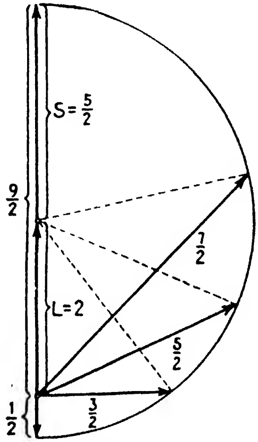


FIG. 86. The vectorial addition of the orbital momentum $\mathbf{L} = 2$ and the resulting spin $\mathbf{S} = \frac{5}{2}$. The five possibilities give the total momentum quantum numbers $J = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}$, and $\frac{9}{2}$ of the atom.

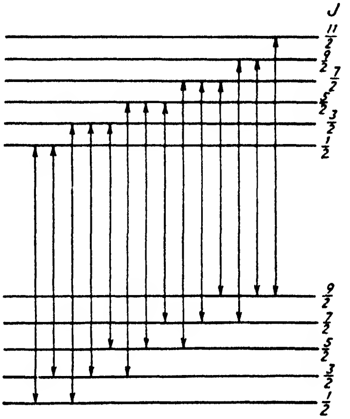
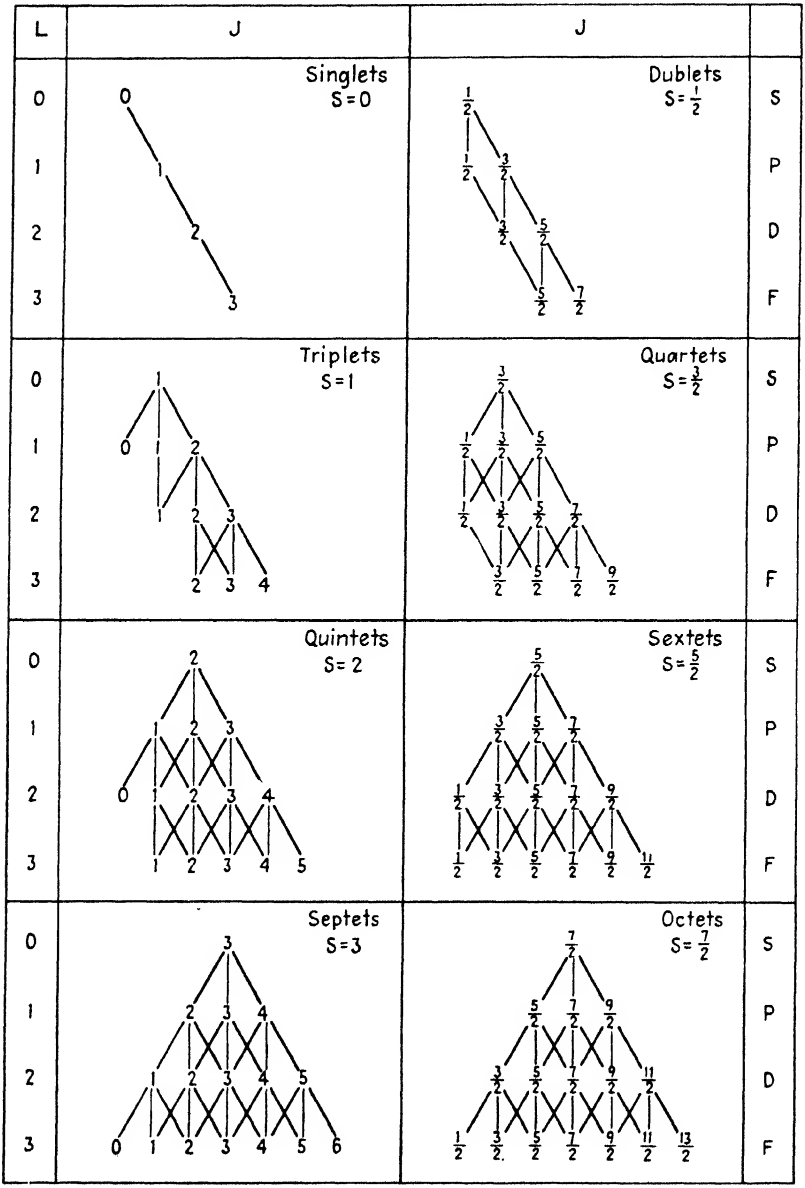


FIG. 87. Term diagram with transitions (schematic) for the vanadium sextet ${}^3F \longrightarrow {}^3D$, shown in Fig. 78.

In Table 7, drawn first by Landé, the possible J values corresponding to the various L and S values, i.e., the term multiplet components, are clearly presented. The transition possibilities, i.e., the number of lines of the associated observed spectral multiplets, are represented by lines.

In summarizing, we can state that even the most complicated phenomena in the spectra of many-electron atoms, such as the existence of

Table 7. Multiplet Table after Landé, Indicating the Possible J Values for Given L and S, and the Possible Transitions between the J Values



several nonintercombining term systems and the occurrence of multiplets, which at first could not be understood, can be explained most satisfactorily by the introduction of the half-integral electron spin and the vectorial combination of electron spin and electron orbital momentum to form the quantized total momentum **J** of the atom.

We have to discuss one more point. From our presentation, because of the spin $\frac{1}{2}$, all single-electron atoms must have a doublet-term system. What about the hydrogen atom which is the most important one-electron atom? Actually, a fine structure of the hydrogen lines (Balmer lines) was found when the spectrum was studied with high resolution. At first, Sommerfeld thought that, because of relativistic effects, this was due to the only approximate coincidence of the S, P, and D terms of the same principal quantum number. On the other hand, the spin theory according to Goudsmit and Uhlenbeck, which requires doublet terms, in the case of one-electron atoms, must also be applicable to the terms of the H atom. This theory leads to a splitting of the two states with principal quantum numbers 2 and 3 as shown in Fig. 88. Transitions between these states account for the fine structure of the Balmer line H_α. Figure 88 shows that the terms split into components of different j quantum numbers, whereas states of the same j but different l, such as ${}^2S_{\frac{1}{2}}$ and ${}^2P_{\frac{1}{2}}$ or ${}^2P_{\frac{3}{2}}$ and ${}^2D_{\frac{3}{2}}$ are, in the first approximation, coincident. The five transitions shown in Fig. 88 (the dotted one is forbidden) result from the application of the selection rule (3-80). This leads to a somewhat different line pattern than was to be expected from the older Sommerfeld theory. Naturally a single-electron ion such as He^+ must show the same behavior. In this case the first line of the Fowler series (see page 112), $\lambda = 4686$ A, is a very sharp line and thus well suited for an investigation. Extremely careful measurements by Hansen on the Balmer line H_α and by Paschen on the He^+ line $\lambda = 4686$ A have established the fine structure of the lines, which was to be expected from Fig. 88. These experiments definitely established the influence of the electron spin on the fine structure of the H atom and the He^+ ion.

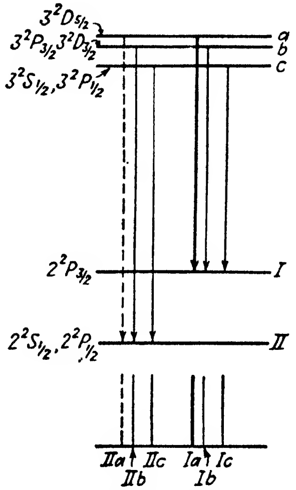


FIG. 88. Term diagram with transitions and resulting spectral line components for the fine structure of the Balmer line H_α (forbidden transition and corresponding line component are dotted).

Very accurate measurements in the centimeter wave region (page 72) have shown, among other results, that the terms $2^2S_{\frac{1}{2}}$ and $2^2P_{\frac{1}{2}}$, shown superimposed in Fig. 88, actually do not coincide, but that the first lies about $\frac{1}{11}$ of the distance I–II (about 0.33 cm^{-1}) higher. This effect is due to the small deviation of the magnetic moment of the electron from the value of Bohr's magneton (page 36). It can be explained only from the most recent development of quantum mechanics (interaction between electron and radiation field).

3-13. The Explanation of the Magnetic Properties of Atoms from the Quantum Theory

Magnetism is one of the properties of matter which can be understood only from the atomic theory. Because it is caused by the orbital revolution and the spin of the electron, we shall devote a section to it at this time. Of the three kinds of magnetism—paramagnetism, diamagnetism, and ferromagnetism—the first two are decidedly properties of single atoms. The last kind is connected with crystal properties and will be treated on page 449 during the discussion of solid state physics. The fact that single iron atoms as well as iron ions in a solution containing an iron compound manifest only paramagnetism is evidence that ferromagnetism is not an atomic property. On the other hand, certain mixed crystals of copper and manganese are ferromagnetic.

In experimental physics two types of magnetic fields are often distinguished: one which is the result of an electric current, and another which seems to have its origin in the magnetic material itself. The results of atomic physics, described in the last section, leave no doubt that the latter kind of magnetic field, in terms of our pictorial model, arises from electric currents set up by electrons rotating in their orbits or by the rotation of electrons about their own axes (spins). Ampère's molecular currents, which were introduced to explain atomic magnetism, thus have found an explanation in the atomic theory. On page 450 we shall show that ferromagnetism is due to a parallel alignment of the spin momenta of all, or nearly all, outermost electrons of the atoms within a crystal domain. However, we must explain dia- and paramagnetism from the structure of the atom itself.

Whether an atom is dia- or paramagnetic depends on the arrangement of its electrons. This can be determined from the spectrum and is to be seen immediately from the term symbols of the ground state (Table 10). For example, in the case of the helium atom we see from its singlet ground state that the two spin momenta of the electrons, and consequently their magnetic moments, are oppositely directed and thus cancel. Because the ground state is an S state, the resultant orbital momentum **L**, and the corresponding magnetic moment, are zero. Consequently, *the helium atom, and all other atoms with a* $^1S_0$ *ground state, can have no magnetic moment.* The same holds for certain diatomic molecules, such as H_2, whose atoms have an S ground state. Although the resultant spin moments of the individual atoms differ from zero (for the H atom, $S = \frac{1}{2}$), they cancel in the molecule, so that the molecule has a singlet ground state.

Thus in all these cases paramagnetism is not possible. It may at first be expected that these atoms and molecules will exhibit no magnetic behavior at all, i.e., we might expect a permeability $\mu = 1$. That their permeability is actually less than one and that they are *diamagnetic* is due to a secondary effect of the external magnetic field which is used to study the magnetic properties of these atoms or molecules. Namely, if we have two electrons with oppositely directed orbital moments which would normally cancel, one of them is accelerated in the magnetic field while the other is slowed down. Since the quantum orbits are not changed, a permanent magnetic moment is thus induced in the atom. The direction of this moment is such that it tries to reduce the generating magnetic field. This effect of reducing an external field, however, is the characteristic property of diamagnetic materials. This inducing action of the external field on the electrons, which causes the diamagnetic behavior of materials without magnetic spin moments, is also present in the case of atoms and molecules with magnetic spin moments, but here the correspondingly small reduction of the magnetic moments is only a small secondary effect. The diamagnetic behavior of atoms with several normally compensating orbital momenta is thus easily explained. That a diamagnetic moment can be induced in atoms such as the He atom with its two 1s electrons by an external magnetic field appears at first to be inexplicable, since we know that the s electrons have orbital momenta $\mathbf{l} = 0$ (page 130). This difficulty will be explained in the next chapter on quantum mechanics, where it will be shown that the s electrons do have orbits or orbital momenta, but that these vanish when averaged over time.

The magnetic moment of paramagnetic atoms is due to the electron rotating in its orbit or to the spin of the electron. The so-called magneto-mechanical parallelism enables us to distinguish between these two contributions to the atomic magnetism. By computing the magnetic moment $\mathfrak{M}_L$ of an electron which has a mechanical orbital momentum $\mathbf{L} = mr^2\omega$ due to its rotation with an angular velocity ω, we find

$$\mathfrak{M}_L = -\frac{e}{2mc}\mathbf{L} \tag{3-82}$$

This relation between the mechanical and magnetic moment can be tested experimentally. Since the law of conservation of the angular momentum must also apply to a system consisting of magnetic atoms, any change of the atoms must cause a corresponding change of the angular momentum, i.e., angular velocity in the opposite direction. This macroscopic change of the angular momentum of a magnetic rod due to any change in its magnetization by an external field is called the Richardson–Einstein–de Haas effect. It can be measured by means of a rotating mirror arrangement as shown in Fig. 89. The opposite effect, the magnetization by rotating the whole system, is called the Barnett effect. The first effect was used to test equation (3-82). *With ferromagnetic materials the effect is large enough so that measurements can be made with high accuracy. The results of the experiment were that the value of the magnetic moment was twice that which was expected from* (3-82). This result, together with the spectroscopic evidence, discussed on page 160, led to the conclusion that the magnetism of ferromagnetic materials is caused by the magnetic moments of electrons, which in turn is due to their eigenrotation (spin). This mechanical spin momentum must have the value $\frac{1}{2}(h/2\pi)$ in order to account for the doubled magnetic moment found in the experiment on ferromagnetic materials. *From the determination of the relation between the mechanical and magnetic moments the distinction can be made whether the magnetic moment is due to the electron rotating in its orbit or to the electron spin.*

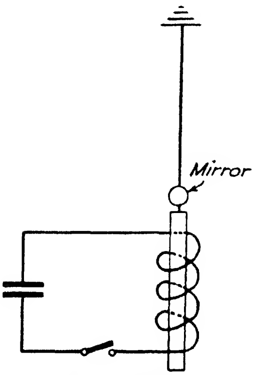


FIG. 89. Schematic arrangement for the Einstein–de Haas effect. Magnetization is caused by discharge of condenser through coil.

In the general case of atoms with orbital momenta **L** and resultant spin momenta **S** of the electron shells, the magnetic moment of the atom depends on the quantum numbers L, S, and J. It can be computed from the formula

$$\mathfrak{M} = \sqrt{J(J+1)}\, g\mu_0 \qquad (3\text{-}83)$$

where μ_0 is the Bohr magneton,

$$\mu_0 = \frac{eh}{4\pi m_e c} \qquad (3\text{-}84)$$

and g the so-called Landé factor,

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \qquad (3\text{-}85)$$

For the case $S = 0$, i.e., pure orbital magnetism, we have $J = L$, and from (3-85), $g = 1$, which corresponds to the normal magnetomechanical relation (3-82) between mechanical and magnetic moment. For pure spin magnetism, on the other hand, we have $L = 0$. Since then $J = S$, Eq. (3-85) gives $g = 2$, which corresponds to the Goudsmit-Uhlenbeck assumption. In all other cases in which both orbital and spin magnetism are present, g has a value different from 1 or 2.

Because of the random temperature motion and random orientation, no magnetism can be noticed, even for atoms with a magnetic moment, without an external field. In an external magnetic field, the orientation of the magnetic atoms is the more complete, the larger the magnetic moments and the field, and the lower the absolute temperature. Because of the space quantization, to be discussed on page 156, the magnetic moment of an atom cannot assume any arbitrary direction relative to the direction of the field. Its components in the direction of the field, $\mathfrak{M}_E$, must be integral multiples of $g\mu_0$,

$$\mathfrak{M}_E = \mathfrak{M} g \mu_0 \tag{3-86}$$

Since the energy difference of the different quantized orientations is generally small compared to the thermal energy kT, only a small fraction of the paramagnetic atoms is oriented in the normally attainable magnetic fields. The system is still far from a saturated state of paramagnetism. The resulting magnetic moment per unit of volume, the magnetization P, under these conditions follows the Curie law,

$$|\mathbf{P}| = \frac{J(J+1)g^2\mu_0^2 N}{3kT} H \tag{3-87}$$

Consequently, P is proportional to the field strength H, and inversely proportional to the absolute temperature T. Thus para- and diamagnetism are explained by the quantum theory.

3-14. Atoms in Electric and Magnetic Fields. Space Quantization and Orientation Quantum Numbers

So far we have learned of three different quantum numbers which characterize the properties of atomic electrons: the principal quantum number n, which approximately indicates the energy; the orbital momentum quantum number l, which, in the Bohr theory, specifies the shape of the orbit; and the spin quantum number which specifies the magnitude of the spin momentum. Spectroscopy has shown that still a fourth quantum number is required to describe completely an atomic electron. This quantum number, which specifies the orientation of one of the angular momenta (**L** or **J**) with respect to an external, or in some cases, internal electric or magnetic field, may be called the orientation quantum number m (M, if it refers to total electron shells). The modification of spectra resulting from an energy change or a splitting of energy states in a magnetic field was discovered by Zeeman in 1896. In 1913, Stark discovered a similar phenomenon with atoms in an electric field. The spectroscopic phenomena caused by these fields, consisting of a splitting or displacement of spectral lines, are called the Zeeman effect and Stark effect, respectively.

In both cases we have a precession about the field direction of the small atomic gyroscopes, characterized by the resultant orbital momentum **J**. In the Stark effect this is due to an existing electric moment or to one induced by polarization. In the Zeeman effect it is due to the magnetic orbital or spin momentum of the atom. The important point here is, that, according to the quantum theory, not every angle between **J** and the field direction is possible, but only those for which the components of **J** in the field are integral or half-integral multiples of $h/2\pi$, depending on whether J itself is integral or half-integral (page 150). This is called space quantization or directional quantization. The difference between the Stark and the Zeeman effect consists in the kind and magnitude of the influence of the electric or magnetic fields on the energy states. *The importance of the Zeeman effect in atomic physics depends, first of all, on the possibility of determining the quantum numbers L, S, and J of the atomic terms from the separation pattern.* The Stark effect was one of the first test examples for the ability of the quantum theory to account for a fairly complicated atomic process. The Stark effect is also of considerable importance in molecular theory, because the molecular electrons in diatomic molecules move in the axial electric field determined by the two positive nuclei, and the theory of the electronic states of molecules (page 359) is based on the behavior of atomic electrons in the Stark effect.

a. Space Quantization and the Stern-Gerlach Experiment

The meaning of space quantization can be most easily understood in the case of the magnetic field. Imagine a small bar magnet so oriented in a magnetic field of strength H that it is at an angle α with the direction of the field (Fig. 90). Then, if its magnetic moment is $\mathfrak{M}$, the magnet has a potential energy

$$U = \mathfrak{M}H \cos \alpha = \mathfrak{M}_H H \tag{3-88}$$

if we designate by $\mathfrak{M}_H$ the component of $\mathfrak{M}$ in direction of H. Consequently, the magnet will oscillate about the direction of H. Furthermore, if it is assumed to rotate about its own axis, it will be deflected perpendicularly to the field as a result of gyroscopic forces. Under these conditions the magnet will precess about the direction of the field with the so-called Larmor frequency, derived in classical physics, which is proportional to the field strength H. As we have already mentioned, not every angle α between the angular momentum $\mathbf{J}$ and direction of the field is possible according to the quantum theory. Only those angles are allowed for which the components $\mathbf{M}$ of $\mathbf{J}$ in the direction of H differ by an integral multiple of $h/2\pi$. M is half-integral or integral, depending upon the value of J. Consequently, an orientation quantum number M is defined by

$$\mathbf{M} = M\frac{h}{2\pi} \tag{3-89}$$

The possible values of M are

$$M = J, J - 1, J - 2, \ldots, -J \tag{3-90}$$

Therefore, there are $2J + 1$ possible values of M.

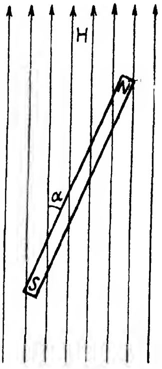


Fig. 90. Energy of a magnetic rod in a magnetic field.

Stern and Gerlach demonstrated this phenomenon of directional quantization by an impressive experiment. They shot a beam of Ag atoms through an *inhomogeneous* magnetic field, Fig. 91, and observed the deviation of the Ag atoms on the photographic plate P. Since the Ag atom with a ground state $^2S_{\frac{1}{2}}$ has a total angular momentum $J = \frac{1}{2}$, the only possible orientations are $M = +\frac{1}{2}$ and $M = -\frac{1}{2}$, with no intermediate positions. In the Stern-Gerlach case, the magnetic field, due to its inhomogeneity, not only produces a quantized orientation of the atomic magnets but also a spatial separation of atoms of both orientations. Without space quantization any orientation of $\mathbf{J}$ with the field would be possible, and the trace of the Ag atoms, impinging on the plate P, would be a broad band. In contrast to this

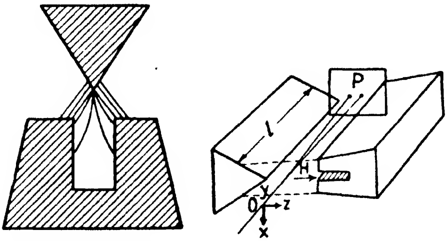


Fig. 91. Schematic representation of the Stern-Gerlach experiment. At the left is a cross section through the pole pieces showing the field lines of the strongly inhomogeneous magnetic field. At the right the separation of the injected atomic beam into two beams by the inhomogeneous magnetic field is shown. P = photographic plate. (*After Briegleb.*)

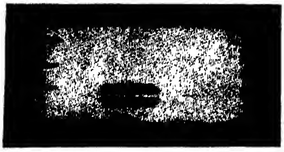


FIG. 92. Separation of a beam of lithium atoms into two distinct beams corresponding to the two possible spin orientations. The arrangement used was that of Fig. 91. (*Photograph by Taylor.*)

classical expectation, two separate traces are found. Space quantization actually caused a splitting of the atomic beam into two beams which correspond to the two possible orientations $M = +\frac{1}{2}$ and $M = -\frac{1}{2}$. Figure 92 is a photograph obtained in a Stern-Gerlach experiment which demonstrates this space quantization.

With unique experimental skill and ingenuity, Rabi and coworkers have developed this comparatively simple Stern-Gerlach experiment into a variety of methods for measuring, with highest precision, the magnetic moments and g values (see Eq. 3-85) of atoms and atomic nuclei.

b. The Normal Zeeman Effect of Singlet Atoms

According to Eq. (3-88), atoms in a magnetic field have different energies depending on their orientation with the direction of this magnetic field. An atomic state characterized by the total angular momentum quantum number J therefore is split up into $2J + 1$ different energy levels in a magnetic field. The corresponding splitting of the spectral lines is called the Zeeman effect. We already know that there are two types of atomic magnetism, the orbital magnetism and the spin magnetism. The behavior of the singlet states is comparatively simple. In this case we deal exclusively with the *orbital magnetic moment* of the atom, which leads to the so-called normal Zeeman effect. On the other hand, the more complex anomalous Zeeman effect, page 160, of the nonsinglet states depends on the magnetic anomaly of the spin magnetism, i.e., on the fact that the magnetic spin moment has twice the normal value compared with the mechanical spin momentum.

First we consider the normal Zeeman effect, in which we are concerned only with the singlet states in which $J \equiv L$. According to the magnetomechanical parallelism, page 153, the magnetic moment $\mathfrak{M}$ is related to the mechanical angular momentum $\mathbf{J}$ by the expression

$$\mathfrak{M} = -\frac{e}{2mc}\mathbf{J} \tag{3-91}$$

The negative sign is due to the negative charge of the electron. Since

$$|\mathbf{J}| = J\frac{h}{2\pi} \tag{3-92}$$

the above equation becomes

$$\mathfrak{M} = -\frac{eh}{4\pi mc}\mathbf{J} \tag{3-93}$$

Thus *the magnetic moment of an atom with* $J = 1$ *is, just as the magnetic moment of an electron, equal to the Bohr magneton*

$$\mu_0 = \frac{eh}{4\pi mc} \tag{3-94}$$

Since the orientation quantum number M represents the quantized component of **J** in the direction of the magnetic field, (3-90), the energy separation of the term components, i.e., their energy displacement from the unmodified term, in a magnetic field H follows from (3-88), (3-93), and (3-94),

$$\Delta E = \mathfrak{M}_H H$$
$$= \frac{eh}{4\pi mc} HM = \mu_0 HM \tag{3-95}$$

Each atomic state is split up into $2J + 1$ equidistant components, whose separation, according to (3-95), is proportional to the magnetic field strength H. The normal Zeeman splitting of neighboring term components ($\Delta M = 1$), computed in wave numbers, is

$$\Delta\bar{\nu}_{\text{norm}} = \frac{\mu_0}{hc} H$$
$$= 4.67 \times 10^{-5} H \quad \text{cm}^{-1} \tag{3-96}$$

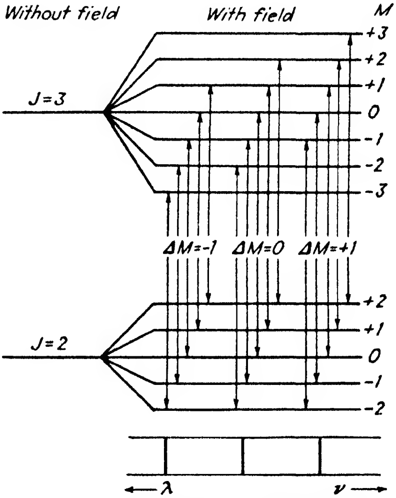


FIG. 93. Term separation, transitions, and resulting "normal Zeeman triplet" in the case of the normal Zeeman effect. The transitions of each of the indicated groups coincide and thus form one component of the observed triplet.

Figure 93 shows the splitting up of two atomic states combining with each other. Since the selection rule which holds for J also applies to the magnetic quantum number M,

$$\Delta M = 0 \text{ or } \pm 1 \tag{3-97}$$

we always find three lines, the normal Zeeman triplet, regardless of the number of term components. This is due to the fact that the energy separation of the components in the upper and lower state has the same value, so that all transitions with the same ΔM coincide in the spectrum. In Fig. 93 these coinciding transitions are collected in groups.

c. *The Anomalous Zeeman Effect and the Paschen-Back Effect of Nonsinglet Atoms*

Empirically, the Zeeman patterns of the nonsinglet states are distinguished by the large number of components and their varying separations. However, Runge showed that the separations are always rational multiples of the normal separation (3-96) (Runge's law). This complexity is due to the combined effects of the orbit and spin magnetism and the magnetic anomaly of the spin. Also in the case of the anomalous Zeeman effect, there is a precession of **J** with the quantized component **M** about the direction of the field. **J** consists, in this case, of the vectorial sum of **L** and **S**, page 147, and this Russell-Saunders coupling is not affected if the magnetic field is not too strong. Because of the magnetic anomaly of the spin magnetism, the components of the magnetic moments associated with the mechanical angular momenta **L** and **S** are

$$\mathfrak{M}_L = -\frac{e}{2mc}\mathbf{L} \qquad (3\text{-}98)$$

and

$$\mathfrak{M}_S = -\frac{e}{mc}\mathbf{S} \qquad (3\text{-}99)$$

Thus there is no factor 2 in the denominator of (3-99). For this reason, as can be seen from Fig. 94, the resultant atomic moment $\mathfrak{M}$, with components $\mathfrak{M}_L$ and $\mathfrak{M}_S$, does not coincide with **J**; its value depends upon L, S, and J.

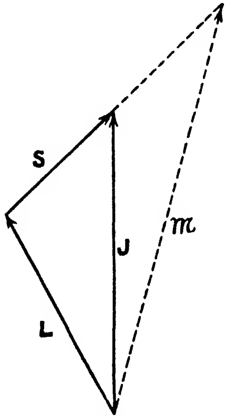


FIG. 94. Vector addition in the case of the anomalous Zeeman effect. As the magnetic moment per unit mechanical spin **S** is twice as large as that per unit **L**, the resulting magnetic moment $\mathfrak{M}$ does not coincide with the resulting mechanical angular momentum **J**.

A nonsinglet atomic state with total momentum quantum number J consequently is split up into $2J + 1$ components distinguished by the various M values. The separation of the components from the undisplaced term now depends on the quantum numbers L, S, and J, and in general, therefore, is not the same for the upper and lower states, so that according to the selection rule (3-97) the Zeeman patterns of the lines consist of many components and are very complicated. The term separation as a function of L, S, and J is expressed by a factor $g(L,S,J)$, so that instead of (3-95), the expression for the spacing of the terms in the anomalous Zeeman effect is

$$\Delta E = \mu_0 g(L,S,J)HM \qquad (3\text{-}100)$$

This g factor is nothing else than the Landé factor of page 154, which was derived first in connection with the Zeeman effect,

$$g(L,S,J) = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \qquad (3\text{-}101)$$

Since g is always a rational number between 1 and 2, the term splitting, according to (3-100), is always a rational multiple of the normal splitting (3-96). Thus Runge's rule is explained theoretically.

The derivation of the Landé factor by Bohr's quantum theory gives

$$g = 1 + \frac{J^2 + S^2 - L^2}{2J^2} \tag{3-102}$$

whereas quantum mechanics, which will be discussed in the next chapter, requires that instead of J the expression $\sqrt{J(J+1)}$ be used. *A point in favor of the new quantum theory and against the old Bohr theory is the fact that the new form of the Landé factor* (3-101) *in the test case of the anomalous Zeeman effect is in exact agreement with the experimental results. Since the determination of the term separation* (3-100) *from the line pattern in the anomalous Zeeman effect depends specifically on the quantum numbers L, S, and J* (3-101), *the anomalous Zeeman effect is a most important aid in determining the quantum numbers of atomic states and, for this reason, is of great importance for theoretical term analysis.* As an example of term splitting in the anomalous Zeeman effect, Fig. 95 shows the terms and the resultant line splitting for the two sodium D lines.

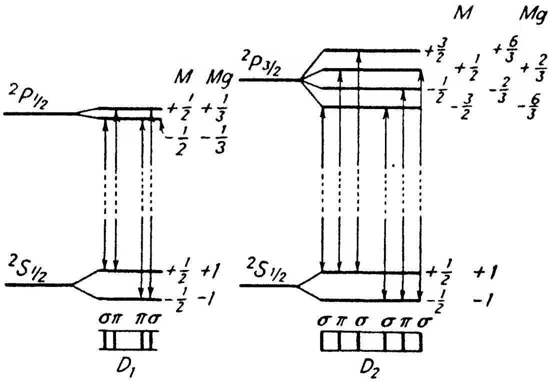


FIG. 95. Term diagrams, transitions, and spectral line components in the case of the anomalous Zeeman effect of the two sodium D lines. (*After Herzberg.*)

If the magnetic field is so strong that the resulting Zeeman splitting (3-100) is as large as the multiplet splitting due to the **L-S** interaction, a new effect found by Paschen and Back occurs. Then the Russell-Saunders coupling between **L** and **S** is uncoupled by the magnetic field, and **L** and **S** will no longer precess about **J**, and **J** about the direction of the field. **L** and **S**, having been uncoupled, will precess independently about the direction of the field, with the quantized components

$\mathbf{M}_L$ and $\mathbf{M}_S$, respectively (Fig. 96). When the Paschen-Back effect has developed completely, the term splitting becomes again an integral multiple of the normal separation (3-95), because M_L, like L, is always an integral number, whereas M_S, like S, may be a half integer. However, because of the magnetic anomaly of the spin, an integral contribution from M_S appears in the formula for the term spacing,

$$\Delta E = \mu_0 H M_L + 2\mu_0 H M_S \qquad (3\text{-}103)$$

Because of the selection rule (3-97) the Paschen-Back effect always produces the normal Zeeman triplet, though with an additional fine structure of each component because of the still existing L-S interaction. Figure 97 shows an example of the term splitting and the resulting spectrum in the case of the Paschen-Back effect. For the transition between the anomalous Zeeman effect and the Paschen-Back effect the line patterns become extremely complicated and, theoretically, rather difficult to treat.

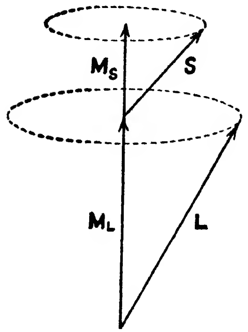


FIG. 96. Vector addition in the case of the Paschen-Back effect.

d. *The Stark Effect*

In the case of the Stark effect the relations are much less simple than in the Zeeman effect, because, in general. an atom does not have an intrinsic electric moment. The Stark effect of the H atom, and of the hydrogen-like atomic states in general, is different from that of all other atomic states, because in the hydrogen case an atomic state of principal quantum number n is $(n^2 - 1)$-fold degenerate, in so far as terms of the same quantum number n but different orbital quantum numbers l normally coincide (page 225). This degeneracy is removed by the electric field, which causes a quantized orientation of the different orbital momenta $\mathbf{l}$ with respect to the direction of the field. As a consequence of this directional quantization, each hydrogen term splits symmetrically into $2n - 1$ term components. The energy differences between these term components, computed from a theory by Schwarzschild and Epstein, are found to be

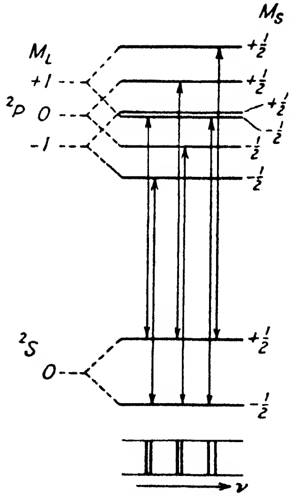


FIG. 97. Term separation, transitions, and line structure for the Paschen-Back effect of a $^2S \rightarrow {}^2P$ transition. (*After Herzberg.*)

in excellent agreement with the results of Stark, and are integral multiples of

$$\Delta E = \frac{3h^2n}{8\pi^2mc}F \tag{3-104}$$

where, in this case, we use F for the electric field strength to distinguish it from the energy E. This result can be interpreted to mean that the external electric field disturbs in a different manner the electrons revolving on orbits of different eccentricities, and thus produces splitting into terms of different orbital quantum numbers l, analogous to the effect of the core of the alkali atoms. The orbital momenta **l** are so oriented in the field that we have a total of $2l + 1$ term components. Because the term separation (3-104) is proportional to the electric field strength, the Stark effect of the hydrogen lines is called linear, in contrast to the more general Stark effect which we shall consider shortly. The term separation in the hydrogen case is, furthermore, according to (3-104), proportional to the principal quantum number n of the atomic states. The explanation of the Stark effect of the Balmer lines with all its details was quite a triumph for the old quantum theory. However, since essential conclusions for atomic physics in general could not be obtained from it, we shall not go into further details.

The Stark effect of the hydrogen atom was a special case because of the high degeneracy of its energy states. In the case of all other atoms, such a degeneracy of states which could be removed by the electric field does not occur, because the energy states associated with the different orbital quantum numbers have already been split up by the field of the atomic core (page 125).

For the non-hydrogen-like atoms the Stark effect depends on a polarization of the atoms by the field, by which a dipole moment proportional to the field is produced. As in the case of the Zeeman effect, the total angular momentum **J** of the atom then precesses about the direction of the field, and it is so directed with respect to the field (space quantization) that its components **M** in the direction of the field are always integral multiples of $h/2\pi$. The energy displacement of the terms compared with the field-free case, as well as the energy splitting of the term components which are characterized by different M's, depend, as in the Zeeman effect, on the product of the dipole moment and field strength. As the former is proportional to the field strength in this general case *the displacement and splitting of term and lines is proportional to the square of the field strength. For this reason, we speak of the quadratic Stark effect.* The details of the quadratic Stark effect are rather complicated because, in addition to the quantum numbers of the split terms, the differences from the neighboring terms play an important role also. In very strong fields, in which Stark effect splitting is large compared to the normal multiplet splitting (or the velocity of precession of **J** about the field direction is large compared to that of **L** and **S** about **J**), **L** and **S** are uncoupled as in the magnetic case and we have an electric analogy of the Paschen-Back effect, page 161.

All these relationships have comparatively little significance for the structure of the atom, because, in contrast to the Zeeman effect, it is not possible to draw direct conclusions from the electrical line splitting about the quantum numbers of the atom. However, the Stark effect is important for the theory of the electronic states of the molecules, page 361. In diatomic molecules, the axis joining the two positive nuclei is the field direction about which the electron orbital momentum precesses. The Stark effect is also of importance for an understanding of the perturbation of atoms by neighboring electrons and ions (line broadening, see page 187). This perturbation can be regarded as an interatomic Stark effect due to the rapidly changing microfields. This Stark-effect broadening plays a considerable role in astrophysics as well as in the physics of gas discharges and of high-temperature plasmas. For example, the density of electrons in a plasma can be determined by measuring the Stark-effect broadening of certain atomic lines.

3-15. The Correspondence Principle and Line Intensities

In our foregoing theoretical discussions we have completely ignored the question of line intensities. The reason for this is obvious. The Bohr theory makes exact statements about the emitted or absorbed frequency or wavelength which corresponds to a quantum transition of an electron, but it tells us nothing about the line intensities. This is in direct contrast to classical wave optics, in which the radiation intensity is proportional to the square of the amplitude of the light wave. These failures are further evidence that the Bohr theory is still a semiclassical theory which only by the artificial introduction of quantum conditions has become suitable for explaining certain atomic phenomena. The quantum-mechanical theory of the atoms, which we shall consider in the next chapter, has all the earmarks of a consistent theory and thus furnishes exact statements about the relation of frequency, intensity, and polarization of spectral lines and continuous spectra.

Bohr himself, with fine physical intuition, patched up the flaws in his theory by establishing a relation with the classical computation of intensity and polarization, called the *correspondence principle*. Even in the case of the absorbed or emitted frequencies of spectral lines there exists a correspondence between the classical and the quantum theory: for high quantum numbers and small changes of the quantum numbers the quantum theoretical frequencies go over into the classical frequencies. We shall show this for the case of the spectral lines of the H atom. According to classical electrodynamics, the frequency of the light radiated by the H atom is equal to the frequency with which the electron rotates in its orbit around the nucleus. Thus from (3-19)

$$\nu_{cl} = \frac{\omega}{2\pi} = \frac{4\pi^2 m e^4}{h^3 n^3} \tag{3-105}$$

From the Bohr theory the frequency emitted in a quantum transition $n_i \longrightarrow n_f$ is

$$\nu_{qu} = \frac{2\pi^2 m e^4}{h^3}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) = \frac{2\pi^2 m e^4}{h^3}\,\frac{n_i^2 - n_f^2}{n_i^2 n_f^2} \tag{3-106}$$

For high quantum numbers and small changes of the quantum number, i.e., when

$$n_i - n_f = \Delta n \ll n_i \tag{3-107}$$

(3-106) becomes

$$\nu_{qu} = \frac{4\pi^2 m e^4}{h^3 n_i^3} \cdot \Delta n \tag{3-108}$$

For the transition between two neighboring quantum states ($\Delta n = 1$) *the classical formula* (3-105) *and quantum formula* (3-108) *become identical. Classically, the transition* $\Delta n = 2$ *corresponds to the first harmonic, etc.*

From this correspondence of the frequencies or wavelengths of the classical and quantum theories, Bohr drew the far-reaching conclusion that intensities and polarization of spectral lines can be calculated from the classical theory; exactly in the case of large quantum numbers, and reasonably correctly in other cases. Experience has confirmed this bold extrapolation. Especially the intensity and polarization of the Stark-effect components of the Balmer lines (page 162), computed by Kramers in 1919 from the correspondence principle, were in such striking agreement with experiment that no further doubt about the usefulness of the correspondence principle was possible. But we must realize that the intensity problem looks entirely different from the classical and quantum-theoretical point of view. Classically, the line intensity is determined by the amplitude of the electromagnetic wave, so that frequency and amplitude are two manifestations of the same phenomenon, the wave. On the other hand, in the quantum theory, the energy, and thus the wavelength, are characteristics of a single radiation process (that is, of a single photon), whereas the intensity of a spectral line depends on the number of photons emitted in the unit volume per unit of time. Consequently, the intensity of a line is independent of individual radiation processes and can be considered a separate statistical problem. From this point of view the correspondence principle appears to be miraculous. Quantum mechanics explains the correspondence principle in clarifying quite generally the relations between classical and quantum physics.

The practical value of the Bohr correspondence principle rests, first of all, on the possibility of using it to derive classically the selection rules for the quantum numbers L, J, and M, which turn out to be in good agreement with experiment. Moreover, the correspondence principle enables us to compute atomic processes classically and then to adjust them, by suitable quantization, to the discontinuous character of the atomic phenomena. In this sense the correspondence principle has some value even in present-day quantum mechanics.

Let us now treat the problem of spectral line intensities in more detail from the quantum-theoretical standpoint. The energy J_ν of frequency ν, radiated by 1 cm³ per second, is equal to the number of atoms in the initial quantum state n_i multiplied by the probability of the transition and the energy of the individual photon, i.e.,

$$J_\nu = N_{n_i} A_{n_i n_f} h\nu \tag{3-109}$$

The transition probability A can be computed exactly from quantum mechanics, page 216, but the correspondence principle enables us to determine classically the value of the expression $A_{n_i n_f} h\nu$. The occupation number of the initial quantum state n_i, i.e., the quantity N_{n_i}, depends on the conditions of excitation. If the excitation is thermal, N_n is given by Boltzman statistics, according to which the occupation of a state of excitation energy E_n at temperature T is given by

$$N_n = N_0 e^{-E_n/kT} \tag{3-110}$$

This formula holds only if there is no degeneracy, i.e., if the energy state E_n is single and does not consist of a superposition of several states of equal energy. Only in this case do we speak of statistical weight 1. On the other hand, an energy state of quantum number J actually consists of $2J + 1$ superimposed energy states, because it can be split up into this number of term components in a magnetic field (Zeeman effect, page 160). The statistical weight of this state, consequently, is $g_J = 2J + 1$. If we call the statistical weight of the ground state g_0 and that of the excited state g_n, we have, instead of (3-110), the more general expression

$$N_n = N_0 \frac{g_n}{g_0} e^{-E_n/kT} \tag{3-111}$$

and the formula (3-109) for the intensity of the radiation becomes

$$J_\nu = N_0 \frac{g_n}{g_0} e^{-E_n/kT} A_{n_i n_f} h\nu \tag{3-112}$$

The probability of an absorption process is proportional to that of an emission, and according to Einstein the probability of a quantum transition $n_i \rightarrow n_f$ under absorption of a photon is

$$B_{n_i n_f} = \frac{c^3}{8\pi h\nu^3} A_{n_i n_f} \tag{3-113}$$

The different quantities in (3-112) usually do not play a role simultaneously so that the relation is essentially simplified. We shall mention a few typical cases. In the case of absorption at moderate temperatures, in general only the ground state is occupied, since the difference of the first excited state E_n from the ground state is always large compared to kT, and thus its occupation is vanishingly small. The statistical weights of the states under consideration in an absorption process can be computed from the quantum numbers and are constant within a series. The intensity distribution in an absorption series, i.e., the decrease of intensity with increasing principal quantum number, consequently reflects directly the dependence of the transition probability on the quantum number of the upper terms. This transition probability thus can be computed from intensity measurements. At a sufficiently high temperature, such as in the atmospheres of the stars, according to page 102, an absorption from excited states is possible, e.g., the absorption of the Balmer series, Fig. 58. The temperature dependence of the absorption of such a series is then given by the Boltzmann factor $e^{-E_n/kT}$. In this case the fraction N_n of atoms in the state under consideration, e.g., of H atoms in the lower state $n = 2$ of the Balmer series, can be determined from intensity measurements, and from these occupation numbers the temperature of the absorbing layer can be calculated. In many nonthermal discharges in which excitation is a result of electron collisions whose energy is large compared to kT, all upper atomic states can be excited with equal probability. In this case the line intensities depend only on the statistical weight of the combining states and on the transition probabilities.

From the dependence of the intensity of lines on the statistical weight (3-112) the following intensity rule for multiplets has been derived: the sum of the intensities of all lines of a multiplet which have the same initial or final state is proportional to the statistical weight $2J + 1$ of that state. We cannot discuss here the consequences of this sum rule for the analysis of multiplets.

3-16. The Explanation of the Periodic Table According to the Atomic Theory

One of the most startling and important results of the Bohr atomic theory was that it led to a satisfactory explanation of the periodic table. Thus the chemical properties of the elements which form the basis of the table (page 16) finally received their atomistic explanation and foundation.

Kossel realized about 1916 that the structure of the electron shells of atoms, deduced from their X-ray spectra (page 117), also accounted for the structure of the periods in the periodic table. This explanation led to Bohr's "construction principle" for the atomic shells. The principle states that the electron shell of each atom is derived from that of the preceding atom by the addition of one more electron. We have already shown on page 116 that the spectroscopic displacement law of Sommerfeld and Kossel can be regarded as direct empirical foundation of this principle of construction. Bohr thus thought that by increasing the nuclear charge in unit steps and by adding one electron with each increase of the positive charge, the atoms of the entire periodic table could be built up from the H atom.

The change of the ionization energies of the atoms with increasing atomic numbers, Fig. 98, offers clear evidence in favor of the shell structure of the electron shells. The second electron, which is built into the second atom, helium, has an ionization energy of 24.6 ev and therefore it is much more closely bound than that of the H atom. Also the second electron must consequently be built into the innermost, $n = 1$, K shell, and both electrons must be closely bound to the nucleus because of its double positive charge. The third electron, which is built into the lithium atom, on the other hand, because of its small ionization energy of only 5.39 ev, must be at a considerably greater distance from the nucleus and, consequently, more loosely bound to it. With lithium the construction of the second shell (L shell) begins. This L shell is completely filled in neon with its ionization potential of 21.5 ev, which is four times as great as that of lithium. We conclude from the smaller ionization potential (5.14 ev) of the next element, sodium, that there the construction of the third shell (M shell) begins. Consequently, it seems that each period of the periodic table corresponds to an electron shell of the atom. The distinct chemical behavior of the extremely active alkali atoms on the one hand, and of the chemically inactive noble gases on the other, thus is understandable from atomic physics: All alkali atoms have one electron outside the closed electron shells, and it is the forces of this electron, the valence electron, which determine the chemical behavior of the atom. The noble gases, however, have closed shells which exert no external forces and thus are not susceptible to chemical binding. We shall return to a more detailed discussion of these problems later (page 409).

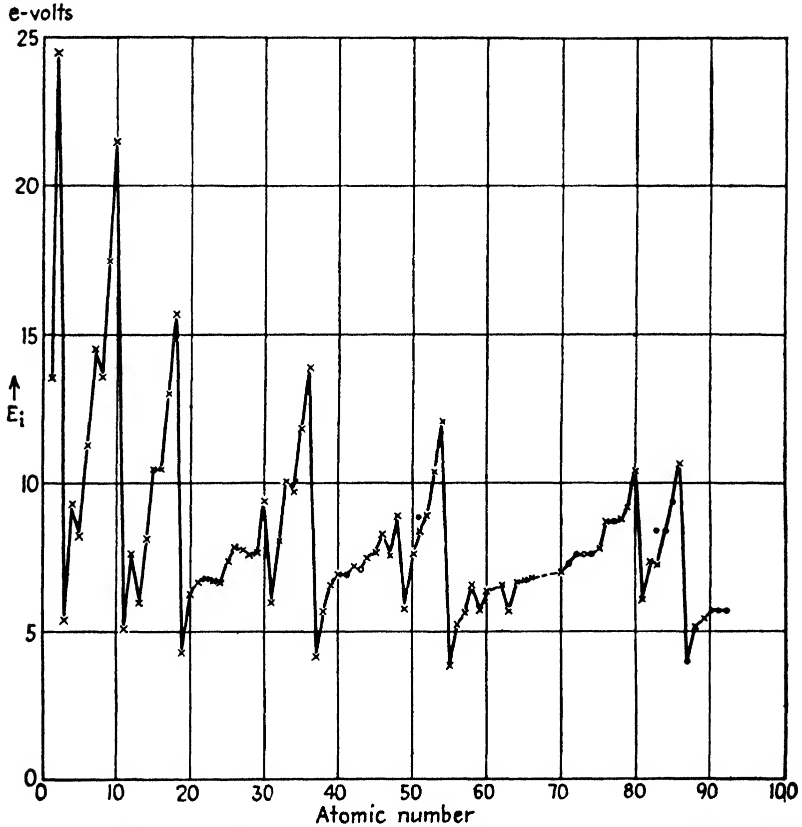


FIG. 98. Ionization potentials of atoms plotted against their atomic numbers. Circles represent values computed or corrected by the author.

This electron-shell structure, deduced from empirical ionization potentials, which will be more firmly and quantitatively established below, has found a sort of theoretical foundation from a principle discovered by Pauli. This principle was not derived theoretically but was deduced from the mass of experimental data. The Pauli principle states: *in no atomic system (atom, molecule, or larger internally bound complex) can there be two electrons which are identical in all four quantum numbers n, l, m, and s.* Using a very mechanistic picture, we might say that we know from experience that there is no "room" in an atom for two electrons which are identical in all properties (quantum numbers). We shall now show that the numbers of elements in the first completely filled periods of the periodic table (2, 8, 18, 32) correspond to the number of unidentical electrons which can, according to the Bohr theory, have the principal quantum numbers $n = 1, 2, 3, 4$. Therefore, they can exist in an atom without violating the Pauli principle. This is shown in Table 8. According to page 130 the maximum orbital momentum of an electron having a principal quantum number n is $l_{\max} = n - 1$. For the different n values, we have, consequently, the different electron possibilities which are shown in column 3 of Table 8. According to page 157, for each l value there are $2l + 1$ different m values (column 4) and each electron characterized by a value of n, l, and m can have a spin oriented in both directions, $s = \pm\frac{1}{2}$ (column 5). For the number of unidentical atomic electrons for each l value we thus have the numbers shown in column 6, and for the number of unidentical electrons belonging to the individual n values (shells) the numbers shown in the last column.

Table 8. Numbers of Unidentical Electrons in the Individual Electron Shells

n	Shell	Possible electrons	Number of different m values	Number of spin orientations	Number of unidentical electrons	
					For a given l	For a given n
1	K	1s	1	2	2	2
2	L	2s	1	2	2	8
		2p	3	2	6	
3	M	3s	1	2	2	18
		3p	3	2	6	
		3d	5	2	10	
4	N	4s	1	2	2	32
		4p	3	2	6	
		4d	5	2	10	
		4f	7	2	14	

It thus follows from the Pauli principle that the number of nonidentical atomic electrons corresponding to the series of principal quantum numbers $n = 1, 2, 3, 4$ in the Bohr-Sommerfeld theory is exactly equal to the number of elements in the closed periods of the periodic table. The construction of the entire periodic table now follows automatically from the construction principle and Table 8 even though, as we shall see, the order of the consecutive electrons is not in complete agreement with that of Table 8.

First of all, both electrons of the normal He atom are 1*s* electrons. Consequently, each has $n = 1$ and $m = l = 0$. Therefore, in order to agree with the Pauli principle, they must have different spin directions. Their spin quantum numbers are $+\frac{1}{2}$ and $-\frac{1}{2}$, as is known from the fact that the ground state of the He atom is a singlet state. There is no room in the innermost shell, the *K* shell, for a third electron, since the *K* shell contains only 1*s* electrons with $l = m = 0$. Moreover, because there are only two possible spin orientations (quantum numbers $+\frac{1}{2}$ and $-\frac{1}{2}$), there can be no third electron which is not identical with one of the others with respect to all four quantum numbers. Consequently, it necessarily follows from the Pauli principle that with the third element lithium a new shell must be formed. In a similar manner, the whole periodic table can be derived from the Pauli principle.

According to Weizel the electron-shell structure, and with it the periodic table, can be more clearly and quantitatively understood by making use of the empirical ionization energies. This also leads to a deeper understanding of the behavior of the atomic electrons.

According to (3-61) the energy necessary to ionize a single-electron atom or ion, if the electron is in the shell with number n, is given by

$$E_i = \frac{Z^2R}{n^2} \tag{3-114}$$

where R must be measured in the same units as E_i.

If we want to apply (3-114) to atoms with more than one electron, we must remember that the whole nuclear charge Z does not act on the valence electron, but that part of the nuclear charge is shielded by the inner electrons. Let Z_{eff} be the effective nuclear charge acting on the outermost electron after taking into account the shielding effect. The empirical ionization energy E_i can then be used to compute Z_{eff},

$$Z_{\text{eff}} = n\sqrt{\frac{E_i}{R}} \tag{3-115}$$

if we know the quantum number n, i.e., the shell number of the external electron. The effective nuclear charge Z_{eff} and the shielding constant

$$s = Z - Z_{\text{eff}} \tag{3-116}$$

have been computed for the first 36 elements of the periodic table from the spectroscopic ionization energies in electron volts, and are presented in Table 9. By assuming that the second helium electron fits into the *K* shell, there follows from its ionization energy of 24.6 ev and from (3-115) an effective nuclear change of 1.35 and a shielding of 0.65 nuclear charges for the second He electron. Since the two electrons are in the same shell, i.e., at the same distance from a charge of $+2e$, it might be

Table 9. Principal Quantum Number of the Emitting Electron, Ionization Potential, Effective Nuclear Charge, and Shielding Constants for the First 36 Elements of the Periodic Table

Element	Atomic number, $Z = N$	Principal quantum number, n	Ionization energy, E_i ev	Effective nuclear charge, $n\sqrt{\frac{E_i}{R}} = Z_{eff}$	Shielding constant, $s = Z - Z_{eff}$	Shielding increment, Δs
H	1	1	13.595	1.00		
He	2	1	24.580	1.35	0.65	0.65
Li	3	2	5.390	1.25	1.75	1.10
Be	4	2	9.320	1.66	2.34	0.59
B	5	2	8.296	1.56	3.44	1.10
C	6	2	11.264	1.82	4.18	0.74
N	7	2	14.54	2.07	4.93	0.75
O	8	2	13.614	2.00	6.00	1.07
F	9	2	17.42	2.26	6.74	0.74
Ne	10	2	21.559	2.52	7.48	0.74
Na	11	3	5.138	1.84	9.16	1.68
Mg	12	3	7.644	2.25	9.75	0.59
Al	13	3	5.984	1.99	11.01	1.26
Si	14	3	8.149	2.32	11.68	0.67
P	15	3	10.43	2.62	12.36	0.68
S	16	3	10.357	2.62	13.38	1.02
Cl	17	3	13.01	2.93	14.06	0.68
A	18	3	15.755	3.23	14.77	0.71
K	19	4	4.339	2.26	16.74	1.97
Ca	20	4	6.111	2.68	17.32	0.58
Sc	21	4	6.56	2.78	18.22	0.90
Ti	22	4	6.83	2.84	19.16	0.94
V	23	4	6.738	2.82	20.18	1.02
Cr	24	4	6.76	2.82	21.18	1.00
Mn	25	4	7.432	2.96	22.04	0.86
Fe	26	4	7.896	3.05	22.95	0.91
Co	27	4	7.86	3.04	23.96	1.01
Ni	28	4	7.633	3.00	25.00	1.04
Cu	29	4	7.723	3.01	25.99	0.99
Zn	30	4	9.391	3.32	26.65	0.69
Ga	31	4	5.97	2.66	28.34	1.66
Ge	32	4	8.13	2.09	28.91	0.57
As	33	4	10.05	3.43	29.57	0.66
Se	34	4	9.750	3.38	30.62	1.05
Br	35	4	11.84	3.73	31.27	0.65
Kr	36	4	13.996	4.06	31.94	0.67

expected that the order of magnitude of the shielding would be $\frac{1}{2}$. Consequently, the theoretical result of 0.65 appears to be reasonable.

If we carry out the computations for the third electron of lithium under the assumption that it is included in the innermost, $n = 1$, shell, we obtain, because of the small ionization potential of lithium, an effective nuclear charge of 0.63. Under these circumstances, the two electrons already in the K shell must shield 2.37 of the 3 positive nuclear charges, a completely senseless conclusion. We therefore conclude that it is not possible for a third electron to be in the K shell. On the other hand, if we compute Z_{eff} for the third electron of lithium under the assumption that it is in the next higher shell, the L shell with $n = 2$, we get from (3-115) the value $Z_{\text{eff}} = 1.25$. In this case each of the two inner electrons of the K shell shields the third electron in the outer shell by 0.87 nuclear charges. This is a very reasonable result. In the case of beryllium we expect a shielding of 0.87 from each of the two inner electrons; and a shielding of 0.65, as for helium, for the third (2s) electron. Together the three electrons should provide a shielding of 2.39 as compared with 2.34 obtained by computing the shielding from (3-115) and (3-116).[1] Consequently, the agreement is excellent and indicates that the fourth electron is also a 2s electron. In the case of the fifth atom, boron, the shielding is increased by 1.10. Thus the fifth electron cannot be a 2s electron, since then it would shield only 0.65 nuclear charges. It must be situated even farther away from the nucleus. Nor can it be an $n = 3$ electron, since for this electron the shielding would be about 3.7 instead of 3.44. So we conclude that the fifth electron of boron is built in in such a way that in the time average it is farther away from the nucleus than a 2s electron, but is closer than a 3s electron. According to the quantum-mechanical results, to be discussed on page 227, this is the expected behavior of a 2p electron. This conclusion, that the first 2p electron appears in boron, is in agreement with the spectroscopic results, that the ground state of boron is a $^2P_{\frac{1}{2}}$ state. The next two electrons, found in the C and N atoms, are also 2p electrons; in agreement with spectroscopic results (increasing multiplicity of the ground state) they are p electrons with their spins in the same direction, but they have different orientation quantum numbers m. Three additional 2p electrons are built into the O, F, and Ne atoms, whose spin is directed oppositely to that of electrons added to the B, C, and N atoms. This follows from the decreasing multiplicity of the

---

[1] As the author has shown, by introducing the shielding increment Δs in the table of the periodic system, very informative regularities are obtained which make it possible to extrapolate unknown Δs values and to use (3-116), (3-115), and (3-114) for determining unknown ionization potentials.

ground state and the jump in the shielding constant in the transition from N to O (due to the addition of the first electron with oppositely directed spin). With neon the second electron shell, the L shell, is completed: there is no possibility of adding another electron of principal quantum number 2 which is not identical with another electron with respect to all four quantum numbers. All together the orbital and spin momenta of the eight electrons compensate each other; neon has the ${}^1S_0$ ground state of the noble gases.

The large change in the shielding by going from neon to the first element of the third period of the periodic table, sodium, indicates that here the construction of a new shell begins. The ground state ${}^2S_{\frac{1}{2}}$ tells us that the first electron included in this shell is a $3s$ electron. The further building up of the shell to the next noble gas, argon, corresponds exactly to the second period. The third shell will contain a second $3s$ electron (of opposite spin) and six $3p$ electrons. In a preliminary way, with argon the third electron shell (M shell) is filled. We say "preliminarily" because we have not yet made use of the fact that there is still room for $3d$ electrons in the $n = 3$ shell. In fact, because of the five different possibilities of orientation (quantum number m) of the orbital momentum, each with two spin directions (Table 8), there may be altogether ten $3d$ electrons. However, we must conclude from the shielding numbers and the ${}^2S_{\frac{1}{2}}$ ground state that, before the ten $3d$ electrons appear, the construction of the fourth quantum shell ($n = 4$) is started with the alkali atom potassium, having one outer $4s$ electron, and Ca with a second $4s$ electron. For the next element, scandium, Sc, the construction does not go on, as it did in the case of the other periods or shells, by including a p electron. This can be seen from the small increase in the shielding. Moreover, the shielding evidence tells us that here an electron must be added inside the outer shell. For this only the $3d$ electrons are available, in agreement with the spectroscopic result that scandium has a ${}^2D_{\frac{3}{2}}$ ground state. Up to nickel seven more $3d$ electrons are added. Then there follows an especially interesting process in the case of copper. By adding the ninth $3d$ electron we get a bivalent Cu atom with two outermost $4s$ electrons, which, except for one missing $3d$ electron, has its three inner shells completed. However, this electron configuration is not stable. One of the two external $4s$ electrons changes over to the free $3d$ position, whereby the number of external valence electrons, and thus the chemical valence of the atom, decreases by unity. This state, the ${}^2S_{\frac{1}{2}}$ state, is the ground state of copper, whose doublet character shows that there is only one external electron, while the S term shows that the number of d electrons must be completed since there is no resultant orbital momentum. However, the Cu state, mentioned first, with the two external 4*s* electrons and the missing 3*d* electron, is so close energetically to the ground state, that it is reached from the ground state under the influence of very small external forces. Consequently, from this atomic theoretical evidence we expect that copper may possibly have two valences. In fact, chemists have long known that copper can be mono- or bivalent (for instance in Cu_2O and CuO). The theory of the periodic table thus easily explains even detail features of the chemical behavior of individual atoms. Referring back to page 116, we want to point out that for such an irregular arrangement of the electron shells the spectroscopic displacement law cannot remain valid. The Cu^+ ion, resulting from the ionization of the copper atom, because of the altered electron arrangement does not have a term diagram and spectrum similar to the preceding atom, nickel. However, this exception to the spectroscopic displacement law is not only understandable from atomic theory but it is a necessary consequence of it.

In the case of zinc the addition of the second 4*s* electron, which had been drawn into the *d* shell in copper, occurs. For the further six elements up to krypton the six 4*p* electrons are added, by which the fourth shell is tentatively closed. There are lacking, however, not only the ten 4*d* electrons, but also the fourteen *f* electrons which should appear for the first time in the $n = 4$ shell, Table 8. With the elements rubidium and strontium, the construction of the $n = 5$ shell begins with two 5*s* electrons, whereupon, as in the preceding period, the further inclusion of the ten 4*d* electrons begins with the third element yttrium. This construction proceeds somewhat differently than in the fourth period, since after the inclusion of the third 4*d* electron in columbium, Cb, one of the external 5*s* electrons is drawn into the *d* shell and thus Cb, like the following elements, is monovalent. After the inclusion of the ninth 4*d* electron, in palladium, the last 5*s* electron is used to fill the *d* shell, as shown by the ground term $^1S_0$ of Pd. Thus the particularly noble, i.e., chemically inactive, character of palladium is explained by the atomic theory as due to the completion of the *d* shell without any outer electrons existing. For the following eight elements from silver to xenon the two 5*s* electrons are again included, and with the inclusion of six 5*p* electrons a tentative completion of the fifth shell is obtained, which is indicated by the noble-gas character of xenon.

In the large sixth period, which begins with the alkali metal cesium and ends with the noble gas radon, Rn, the behavior is especially complicated. However, even here the explanation from the atomic theory is in excellent agreement with the chemical behavior. In the case of the first elements Cs and Ba the structure of the sixth shell begins with the inclusion of the 6*s* electrons. With the next element lanthanum, as for the corresponding place in the fifth period, the supplementary addition of the 5*d* electrons begins. However, with the next element cerium, there begins a series of elements which are very much alike chemically and which therefore are difficult to isolate, the *rare earths*. They are built up by the supplementary inclusion of the 4*f* electrons which we mentioned above. Since these are built into the $n = 4$ shell, beyond which the $n = 5$ and $n = 6$ shells are already partly completed, they do not have much effect on the chemical behavior of the rare-earth atoms. The chemical similarity and the difficulty of separation could not be explained from a purely chemical point of view. However, this explanation follows as an evident consequence from the electron structure. The 4*f* shell is completed in ytterbium, and in the following elements up to platinum the 5*d* electrons are included along with some of the external 6*s* electrons. From gold to radon the two 6*s* electrons and six 6*p* electrons are built in as usual. For the last elements of the periodic table, the construction follows that of the preceding period. The construction of the $n = 7$ shell begins with the two 7*s* electrons. Then in actinium the first 6*d* electron is included, and with the last seven elements of the periodic system, as in the case of the rare earths, inner *f* electrons, in this case 5*f* electrons, are included. The last elements thorium, protactinium, and uranium, as well as the "trans-uranium" elements discovered in the last few years, neptunium, plutonium, americium, and curium correspond throughout to the rare earths. The rare earths are often called "lanthanides" from their preceding element. For the same reason the elements from thorium through curium are called "actinides."

The step-like completion of the atomic electron shells and, correspondingly, of the periodic table has been represented in a pictorial diagram by Schultze. This diagram, corrected according to our present knowledge, is shown in Fig. 99. The most important data on atoms, the atomic number, the term symbol of the ground state, and the arrangement of the electrons in the different shells, are shown in Table 10.

Let us consider now what has been achieved and what remains to be done. The fundamental idea was that the periodic structure of the periodic table corresponds to the electron-shell structure of the atoms, and can be derived from the construction principle (which itself is a consequence of the spectroscopic displacement law) by consideration of the Pauli principle. On this basis, we could easily explain the periodic table and the chemical properties of atoms. From this explanation there necessarily follow, without any special assumptions, not only

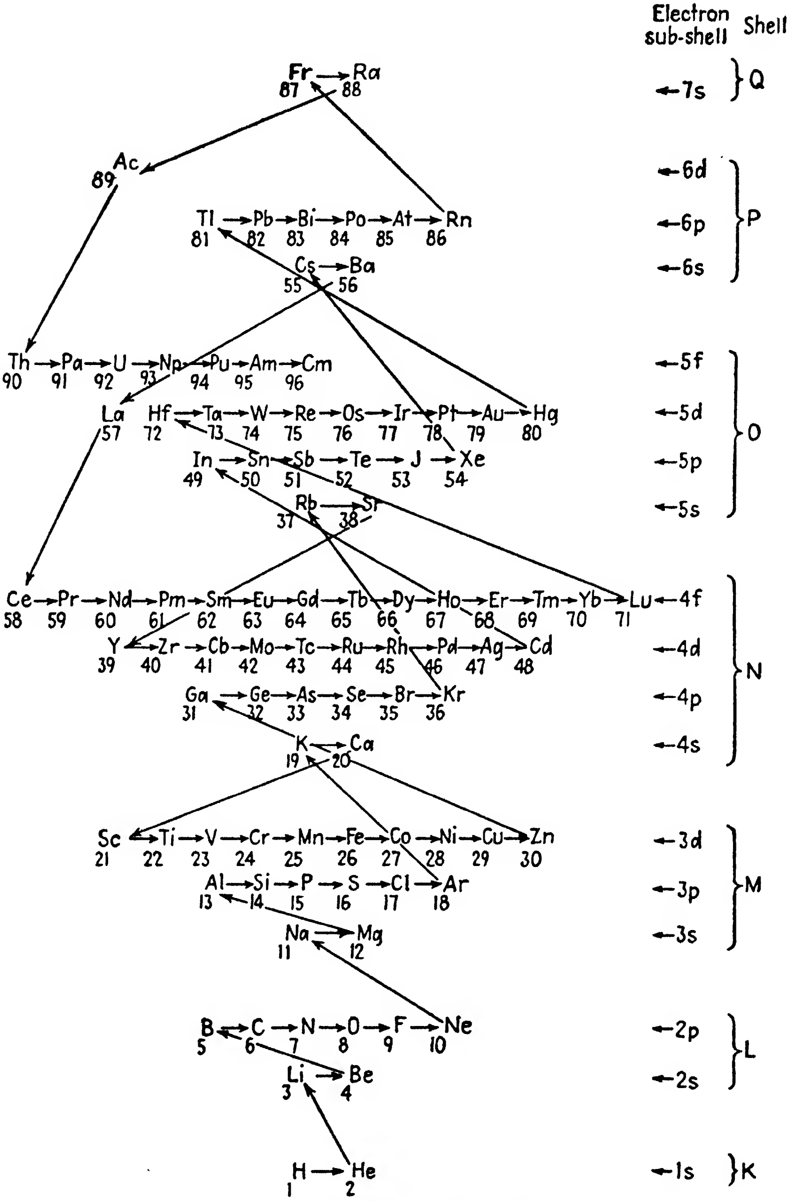


FIG. 99. Electron arrangement in the shells of the 96 atoms of the periodic system. (*After Schultze, with corrections by the author.*)

Table 10. Ground States and Electron Configurations of the Elements

Z			K	L		M			N				O			P			Q
			1s	2s	2p	3s	3p	3d	4s	4p	4d	4f	5s	5p	5d	6s	6p	6d	7s
1	H	$^2S_{\frac{1}{2}}$	1																
2	He	$^1S_0$	2																
3	Li	$^2S_{\frac{1}{2}}$	2	1															
4	Be	$^1S_0$	2	2															
5	B	$^2P_{\frac{1}{2}}$	2	2	1														
6	C	$^3P_0$	2	2	2														
7	N	$^4S_{\frac{3}{2}}$	2	2	3														
8	O	$^3P_2$	2	2	4														
9	F	$^2P_{\frac{3}{2}}$	2	2	5														
10	Ne	$^1S_0$	2	2	6														
11	Na	$^2S_{\frac{1}{2}}$	2	2	6	1													
12	Mg	$^1S_0$	2	2	6	2													
13	Al	$^2P_{\frac{1}{2}}$	2	2	6	2	1												
14	Si	$^3P_0$	2	2	6	2	2												
15	P	$^4S_{\frac{3}{2}}$	2	2	6	2	3												
16	S	$^3P_2$	2	2	6	2	4												
17	Cl	$^2P_{\frac{3}{2}}$	2	2	6	2	5												
18	Ar	$^1S_0$	2	2	6	2	6												
19	K	$^2S_{\frac{1}{2}}$	2	2	6	2	6		1										
20	Ca	$^1S_0$	2	2	6	2	6		2										
21	Sc	$^2D_{\frac{3}{2}}$	2	2	6	2	6	1	2										
22	Ti	$^3F_2$	2	2	6	2	6	2	2										
23	V	$^4F_{\frac{3}{2}}$	2	2	6	2	6	3	2										
24	Cr	$^7S_3$	2	2	6	2	6	5	1										
25	Mn	$^6S_{\frac{3}{2}}$	2	2	6	2	6	5	2										
26	Fe	$^5D_4$	2	2	6	2	6	6	2										
27	Co	$^4F_{\frac{1}{2}}$	2	2	6	2	6	7	2										
28	Ni	$^3F_4$	2	2	6	2	6	8	2										
29	Cu	$^2S_{\frac{1}{2}}$	2	2	6	2	6	10	1										
30	Zn	$^1S_0$	2	2	6	2	6	10	2										
31	Ga	$^2P_{\frac{1}{2}}$	2	2	6	2	6	10	2	1									
32	Ge	$^3P_0$	2	2	6	2	6	10	2	2									
33	As	$^4S_{\frac{3}{2}}$	2	2	6	2	6	10	2	3									
34	Se	$^3P_2$	2	2	6	2	6	10	2	4									
35	Br	$^2P_{\frac{3}{2}}$	2	2	6	2	6	10	2	5									
36	Kr	$^1S_0$	2	2	6	2	6	10	2	6									
37	Rb	$^2S_{\frac{1}{2}}$	2	2	6	2	6	10	2	6			1						
38	Sr	$^1S_0$	2	2	6	2	6	10	2	6			2						
39	Y	$^2D_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	1		2						
40	Zr	$^3F_2$	2	2	6	2	6	10	2	6	2		2						
41	Cb	$^6D_{\frac{1}{2}}$	2	2	6	2	6	10	2	6	4		1						
42	Mo	$^7S_3$	2	2	6	2	6	10	2	6	5		1						
43	Tc	($^6S_{\frac{3}{2}}$)	2	2	6	2	6	10	2	6	(5)		(2)	?					
44	Ru	$^5F_5$	2	2	6	2	6	10	2	6	7		1						
45	Rh	$^4F_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	8		1						
46	Pd	$^1S_0$	2	2	6	2	6	10	2	6	10								
47	Ag	$^2S_{\frac{1}{2}}$	2	2	6	2	6	10	2	6	10		1						
48	Cd	$^1S_0$	2	2	6	2	6	10	2	6	10		2						

Table 10. (Continued)

Z			K	L		M			N				O				P			Q
			1s	2s	2p	3s	3p	3d	4s	4p	4d	4f	5s	5p	5d	5f	6s	6p	6d	7s
49	In	${}^2P_{\frac{1}{2}}$	2	2	6	2	6	10	2	6	10		2	1						
50	Sn	${}^3P_0$	2	2	6	2	6	10	2	6	10		2	2						
51	Sb	${}^4S_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	10		2	3						
52	Te	${}^3P_2$	2	2	6	2	6	10	2	6	10		2	4						
53	I	${}^2P_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	10		2	5						
54	Xe	${}^1S_0$	2	2	6	2	6	10	2	6	10		2	6						
55	Cs	${}^2S_{\frac{1}{2}}$	2	2	6	2	6	10	2	6	10		2	6			1			
56	Ba	${}^1S_0$	2	2	6	2	6	10	2	6	10		2	6			2			
57	La	${}^2D_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	10		2	6	1		2			
58	Ce	$({}^3H_4)$	2	2	6	2	6	10	2	6	10	2	2	6			2	?		
59	Pr	—	2	2	6	2	6	10	2	6	10	3	2	6			2	?		
60	Nd	—	2	2	6	2	6	10	2	6	10	4	2	6			2	?		
61	Pm	—	2	2	6	2	6	10	2	6	10	5	2	6			2	?		
62	Sm	${}^7F_0$	2	2	6	2	6	10	2	6	10	6	2	6			2			
63	Eu	${}^8S_{\frac{7}{2}}$	2	2	6	2	6	10	2	6	10	7	2	6			2			
64	Gd	${}^9D$	2	2	6	2	6	10	2	6	10	7	2	6	1		2			
65	Tb	—	2	2	6	2	6	10	2	6	10	8	2	6	1		2	?		
66	Dy	—	2	2	6	2	6	10	2	6	10	9	2	6	1		2	?		
67	Ho	—	2	2	6	2	6	10	2	6	10	10	2	6	1		2	?		
68	Er	—	2	2	6	2	6	10	2	6	10	11	2	6	1		2	?		
69	Tm	${}^2F_{\frac{7}{2}}$	2	2	6	2	6	10	2	6	10	13	2	6			2			
70	Yb	${}^1S_0$	2	2	6	2	6	10	2	6	10	14	2	6			2			
71	Lu	${}^2D_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	10	14	2	6	1		2			
72	Hf	${}^3F_2$	2	2	6	2	6	10	2	6	10	14	2	6	2		2			
73	Ta	${}^4F_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	10	14	2	6	3		2			
74	W	${}^5D_0$	2	2	6	2	6	10	2	6	10	14	2	6	4		2			
75	Re	${}^6S_{\frac{5}{2}}$	2	2	6	2	6	10	2	6	10	14	2	6	5		2			
76	Os	${}^5D_4$	2	2	6	2	6	10	2	6	10	14	2	6	6		2			
77	Ir	${}^4F$	2	2	6	2	6	10	2	6	10	14	2	6	7		2			
78	Pt	$({}^3D)$	2	2	6	2	6	10	2	6	10	14	2	6	9		1	?		
79	Au	${}^2S_{\frac{1}{2}}$	2	2	6	2	6	10	2	6	10	14	2	6	10		1			
80	Hg	${}^1S_0$	2	2	6	2	6	10	2	6	10	14	2	6	10		2			
81	Tl	${}^2P_{\frac{1}{2}}$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	1		
82	Pb	${}^3P_0$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	2		
83	Bi	${}^4S_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	3		
84	Po	${}^3P_2$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	4		
85	At	${}^2P_{\frac{3}{2}}$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	5		
86	Rn	${}^1S_0$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6		
87	Fr	${}^2S_{\frac{1}{2}}$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6		1
88	Ra	${}^1S_0$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6		2
89	Ac	$({}^2D_{\frac{3}{2}})$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6	1	2 ?
90	Th	$({}^3F_2)$	2	2	6	2	6	10	2	6	10	14	2	6	10		2	6	2	2 ?
91	Pa	$({}^4F_{\frac{3}{2}})$	2	2	6	2	6	10	2	6	10	14	2	6	10	2	2	6	1	2 ?
92	U	$({}^5D_0)$	2	2	6	2	6	10	2	6	10	14	2	6	10	3	2	6	1	2 ?
93	Np	?	2	2	6	2	6	10	2	6	10	14	2	6	10	4	2	6	1	2 ?
94	Pu	?	2	2	6	2	6	10	2	6	10	14	2	6	10	5	2	6	1	2 ?
95	Am	?	2	2	6	2	6	10	2	6	10	14	2	6	10	7	2	6		2 ?
96	Cm	?	2	2	6	2	6	10	2	6	10	14	2	6	10	7	2	6	1	2 ?

the general rules for the chemical behavior of the elements, but also the finer features such as the different valences of copper and the special positions occupied by the noble metals palladium and platinum, as well as the particular properties of the rare earths. This whole explanation is based on the empirically measured ionization potentials and the spectroscopically determined properties of the ground states of the atoms (multiplicity and L values). This explanation can be derived from the Pauli principle without the aid of the empirical data, but not in all details. This principle was derived from experience, and up to the present time it has not been possible to derive it as a necessary consequence of the atomic theory.

The success of the Bohr theory in explaining the periodic table is so undoubtedly great that it may be regarded as its crowning achievement. Nevertheless, our discussion shows very clearly what has yet to be done. The Bohr theory provides an *explanation* of the periodic table; but it is *not a consistent and rigorous theory*. Such a more general theory would have to explain why each consecutive electron takes *just that* and no other place, why after the first $5d$ electron the $4f$ electrons follow, etc. This complete theory of the periodic table, which by necessity would contain a derivation of the Pauli principle, is still lacking. However, by applying his quantum statistics, page 237, to the electrons of the atomic core, Fermi has been able to show that the arrangement of Table 10 is the most stable one of all possible arrangements. On the other hand, the question which we left open in our presentation, why the periodic table breaks off with the ninety-second element, or now with the transuranium elements which have been recently discovered, has in the meantime been answered. The reason for this breaking off is not due, as was believed for a long time, to an instability of the electron shells beginning with uranium. It is due to the acutely increasing instability of the heaviest atomic nuclei. We shall consider this when we discuss the subject of nuclear physics on page 248.

3-17. The Hyperfine Structure of Atomic Lines. The Isotope Effect and the Influence of the Nuclear Spin

The splitting of the atomic terms and the spectral lines into multiplets was discussed in Sec. 3-12, page 148, as being due to the quantized coupling of the orbital momentum and spin momentum of the electrons. It is generally called the *fine structure of spectral lines*, although for heavy atoms this splitting of the terms is so great that the fine-structure components become widely separated lines. In contrast to this fine structure of the spectra, which depends entirely on processes within the electron shells, we have what is called the *hyperfine structure of* *spectral lines* which is due to the interaction of the electron shells with the *nucleus.* The hyperfine structure can be measured only with very highly resolving interference spectroscopes, or by the methods of the new microwave spectroscopy (page 72).

This hyperfine structure can have two fundamentally different origins. In the first place, the atom under consideration may have several isotopes (see Table 3, page 53). The effect of the different masses and different structures of the isotopic nuclei on the electron shells then causes small wavelength differences of the spectral lines belonging to the different isotopes, so that under high resolution what appear to be single lines actually are revealed as very closely spaced lines belonging to the different isotopes. In the second case, most nuclei, like the electron itself, have a mechanical spin momentum, the nuclear spin **I**, and a small magnetic moment (see page 268). It is the interaction of this nuclear moment with the resulting magnetic moment of the electron shell that causes term and line splitting, i.e., the hyperfine structure of spectral lines. The overlapping of both effects naturally complicates the explanation of the hyperfine structure.

We shall consider at first only the isotope effect. We saw when we discussed the H atom, page 98, that as a result of the motion of the nucleus and the electron about a common center of mass, Rydberg's constant R and thus the wavelengths of the spectral lines depend on the mass of the nucleus. If we replace the nucleus of the H atom, the proton, by that of the heavy H isotope, the deuteron, which consists of one proton and one neutron, we get from Eq. (3-26) a change in the wavelength of the first Balmer line of 1.79 A. Urey demonstrated the existence of the hydrogen isotope by measuring this wavelength difference in 1932. In the case of heavy atoms with small mass differences between their isotopes, the λ difference rapidly becomes so small that it falls in the region of hyperfine structure. Small differences of this isotope displacement, depending on the motion of the nucleus, occur also as a result of the rotation of the external electrons of the many-electron atoms in the same or opposite directions. We shall not consider the details of the effect. For the heavy atoms, on the other hand, there is a second isotopic effect which is worthy of notice. For higher masses it exceeds the effect of the mutual motion and is called the *volume effect.* By addition of one or more neutrons to the nucleus, the spatial arrangement of the protons, which account for the binding of the electron shells to the nucleus, is changed. Consequently, since in the case of large nuclei the positive charge can no longer be regarded as a point charge, the energy of the electrostatic field between the nucleus and the electron shells which determines the energy state of the atom, depends upon the arrangement of protons and neutrons in the nucleus. It varies from isotope to isotope, and this results in an isotopic line shift. We shall not go into any details, but simply state the result that the number, masses, and relative abundances of the isotopes of an element can be determined by measuring the number, separation, and intensities of the hyperfine-structure components resulting from the isotope effect. Such optical spectroscopic investigations are an important supplement to mass-spectroscopic research on isotopes. This was discussed on page 46. We shall discuss additional spectroscopic effects of isotopes on molecular band spectra in Chap. 6, page 398.

The actual hyperfine structure of terms and lines of individual isotopes, i.e., after elimination of the eventual isotope effect, is treated theoretically, just as was the multiplet structure in Sec. 3-12. Like the electron, the nucleus also must have a spin moment **I** with a corresponding quantum number I,

$$|\mathbf{I}| = I\frac{h}{2\pi} \tag{3-117}$$

However, in contrast to the electron with a constant spin $\frac{1}{2}$, the spin of the nucleus can have different values, because the nucleus consists of a number of protons and neutrons. According to the measurements given on page 268, I can be between 0 and $\frac{9}{2}$. In addition to the spin momentum there is also a magnetic moment which, according to page 268, has the order of magnitude

$$\mu_n = \frac{eh}{4\pi Mc} \tag{3-118}$$

and is called the nuclear magneton. It can be deduced from the Bohr magneton [(3-94), page 159], by substituting the proton mass M for the mass of the electron m. The value of the nuclear magneton is thus 1/1,837 of the Bohr magneton or the magnetic moment of the electron. Because the value of the nuclear magnetic moment is so small, the interaction between the nucleus and electron shell is also small, and this is the reason for the very small separation of the hyperfine-structure components of the atomic terms.

This nuclear spin **I** is space-quantized in the magnetic field of the electron shells, and vectorially combined with the total momentum **J** of the electron shells forms the quantized total angular momentum **F** of the whole atom. Therefore, the number of hyperfine-structure terms corresponding to this quantum number is

$$2J + 1 \qquad \text{or} \qquad 2I + 1 \tag{3-119}$$

depending on whether $I > J$ or $J > I$, respectively. The *number of the hyperfine-structure terms* thus depends, as in the case of the normal fine structure, on the values of the angular momentum quantum numbers. The *spacing*, on the other hand, depends on the magnitude

$$A = \frac{\mu_I H_n}{IJ} \qquad (3\text{-}120)$$

that is, on the value of the nuclear magnetic moment μ_I [which according to (3-93) depends on the nuclear spin I] and the field strength H_n of the magnetic field set up by the electron shells at the location of the nucleus.

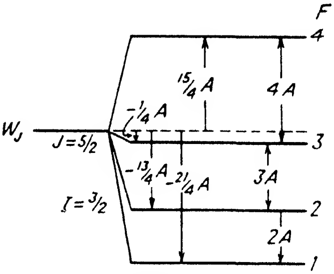


FIG. 100. Hyperfine-structure term multiplet with quantum numbers and term distances. (*After Kopfermann.*)

As an example, we show in Fig. 100 a hyperfine-structure term multiplet for $J = \frac{5}{2}$ and $I = \frac{3}{2}$ whose combination gives the values 1, 2, 3, and 4 for the total angular momentum F of the atom. We cannot go into the computation of the actual spacings. The same selection rule

$$\Delta F = 0 \text{ or } \pm 1 \qquad (3\text{-}121)$$

that applies to the quantum number J is valid for transitions between hyperfine-structure term components.

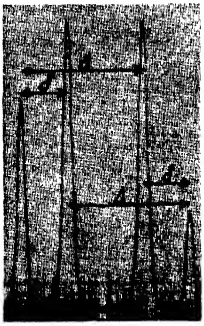


FIG. 101. Densitometer record of the hyperfine-structure components of the Bi line 4122 A. (*After Zeeman, Back, and Goudsmit.*)

Figure 101 shows as an actual example the densitometer curve of the hyperfine structure of the bismuth line $\lambda = 4122$ A which results from the term diagram and transitions shown in Fig. 102. The term difference of the two upper states, δ, and the two lower ones, Δ, occurs twice in Fig. 101 where it can be seen at a glance (compare the general remarks on the empirical term analysis, page 83).

The influence on the Hfs terms of an external magnetic field can be used to prove the accuracy of a hyperfine-structure term analysis. All Hfs terms are split up in a magnetic field into $2F + 1$ components as a result of space quantization. As in the case of ordinary terms (page 161), we distinguish in the case of hyperfine structure between the Zeeman effect and the Paschen-Back effect according to whether the spacing due to the external magnetic field is small or large compared with the natural spacing of the term components due to the internal fields.

The measurement of the number, orientation, and separation of the hyperfine-structure components makes it possible to determine the nuclear spin I and the sign of the magnetic moment μ_I. In favorable cases μ_I itself can be determined. By using the deviation of the measured hyperfine-structure components from those computed under the assumption that the nuclear charge distribution is spherically symmetric, it is possible, according to Schüler, to compute an eventual electric quadrupole moment. Thus, hyperfine-structure experiments have become an important aid to nuclear physics. Recently, they have been carried out to an increasing extent with the methods of microwave spectroscopy (page 72) which are especially well suited for the measurement of such small wavelength differences.

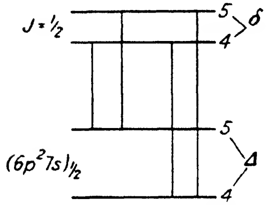


FIG. 102. Hyperfine-structure term diagram of the Bi line 4122 A explaining Fig. 101.

3-18. The Natural Width of Spectral Lines and the Influence of Internal and External Perturbations

Our discussion of the relation of line spectra to atomic structure and to atomic processes would not be complete if we did not say a few words about the interesting question of the width of spectral lines. All spectral lines have a "natural" width which is determined by atomic processes and is independent of the resolution of the spectroscopic equipment. If, namely, the spectral line is not emitted or absorbed by an atom at rest, but by atoms in thermal motion, or by atoms exchanging energy with their environment by kinetic collisions or by the action of an electric field due to electrons or ions, then the effect of these factors can be observed by the change in the width and intensity distribution of the spectral lines. Experiments on these variations permit conclusions to be drawn on the state of motion of the atom and the effect of its environment. These relations are of great importance for detailed studies in electric gas discharges as well as for astrophysicists. Research on the physical conditions in the various layers of the sun's atmosphere (pressure, temperature, particle density) depends to a great extent on the theory of line broadening.

It can be shown classically that spectral lines have a certain natural width, because the radiation consists of individual wave trains of finite length. This is identical with the concept that we have, in the classical picture, an emission of damped waves by Hertzian oscillators. Then a Fourier analysis of the vibration process results in not only *one* frequency but a large number of neighboring frequencies occurring with decreasing intensities. This means that a spectral line actually is a more or less broad band of frequencies or wavelengths, as indicated in Fig. 103. In speaking of the width of spectral lines, we use the expression "half width" when we mean the width at half the maximum intensity. Classical computations give for this natural width of a spectral line the quantity

$$\gamma = 1.19 \times 10^{-4} \text{ A} \tag{3-122}$$

independent of the wavelength. Its value is so small that it is not easy to measure the natural width of a line, and thus, in general, the normal lines of undisturbed atoms appear to be sharp.

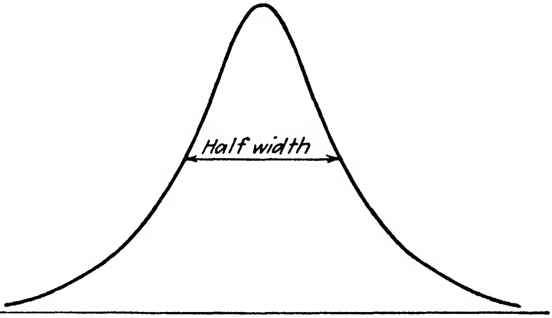


FIG. 103. Contour of a spectral line with half width indicated.

In quantum mechanics the line width is determined by the width of the two combining energy levels. It follows from the uncertainty principle, which we shall consider in the next chapter, page 200, that the energy level of a nonstable atomic state must have a finite width, since the uncertainty principle states that the energy of a state is the less accurately determined, the shorter its mean life τ, or expressed in a formula

$$\Delta E \sim \frac{h}{\tau} \tag{3-123}$$

Here ΔE is the uncertainty in determining the energy, i.e., the width of the energy level under consideration. For resonance lines it can be shown that the formula for the line width which follows from (3-123) is identical with (3-122) which was arrived at by considering damped classical radiation if we use for τ in (3-123) the average life (10^{-8} sec) of an excited atom. However, it is a consequence of the quantum-mechanical formula (3-123) that, in contrast to the classical formula, *the width of the energy level and thus the spectral line not only depend on the mean life with respect to radiation, but also can be limited by processes which do not produce radiation.*

In addition to collisions with neighboring particles (collision broadening) and similar disturbances, there are two nonradiative processes which can limit the mean life of an atomic state. Because of their importance, we shall mention them briefly. They are autoionization and the ionization of an atom in a strong electric field. Autoionization is possible, if, as a result of double excitation (page 136), a stationary atomic state exists above the normal ionization limit. If certain selection rules are fulfilled, then, according to Fig. 104, a transition from the doubly excited state to a lower state accompanied by radiation, or a nonradiating transition to a neighboring continuous energy region, can occur. The latter means that the second excited electron goes into the ground state and the first one, using the released energy, leaves the atom entirely and is ionized. On certain conditions, *an ionization of an atom from a doubly excited state can thus occur instead of an emission of a photon. Therefore, we call this process autoionization.* Another term which is used often is preionization, which means that ionization occurs before the ionization limit of the particular series is reached. The mean life of the excited stationary state is, in this case, limited not only by the emission of radiation, but under certain conditions with higher probability by a nonradiative transition into the continuum. According to (3-123), this shortened mean life corresponds to an increased width of the atomic state and, if the transition occurs to a lower state, a broader line. The width of the lines, which is sometimes very large, can be experimentally measured and serves as evidence for the existence of preionization and as a measure for its probability.

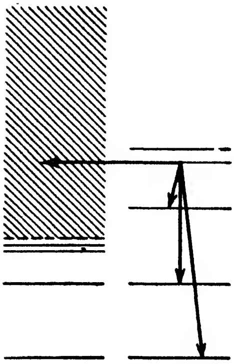


FIG. 104. Term diagram and transition possibilities in the case of autoionization of an atom.

A similar case is the ionization in a strong external electric field. If an H atom is brought into a strong homogeneous electric field E, its potential $-eEx$ overlaps that of the Coulomb field of the nucleus $-e^2/r$, so that the potential of the H atom is

$$U(r) = -\frac{e^2}{r} - eEx \tag{3-124}$$

In Fig. 105 the Coulomb field is plotted with the superimposed homogeneous field shown as a straight dotted line; by adding both fields, we get the potential curve with a maximum shown as a solid line. It can be seen from Fig. 105 that the stationary states of the H atom, lying above the maximum of the potential curve, can no longer be stable. The electron which is normally bound to the nucleus is torn from the atom by the external field, i.e., the atom is ionized. According to an important effect which will be considered in quantum mechanics (tunnel effect, page 234) there is a certain probability that even electrons of somewhat lower energy states will be ionized: if the potential wall is not too high or broad, the electron can penetrate it by means of the tunnel effect. For a stationary state just below the potential maximum there is, thus, in addition to the possibility of jumping to a lower state with the emission of radiation, the possibility of passing through the potential wall without radiating and becoming a free electron. This process tends to further reduce the mean life τ of a stationary state and to increase its width and that of the corresponding line. As a consequence of the processes which have been described, it can be seen that even before the ionization limit is reached, the last series lines disappear into the continuum, and the last lines which can still be recognized show a noticeably increased width because of the possibility of nonradiative ionization from their upper states. Figure 61 shows this effect, which is called "overlapping of the continuum beyond the series limit," in the spectrum of a thallium discharge very clearly. The electric field, which is inhomogeneous with respect to time and space, originates in this case from electrons and ions of the discharge.

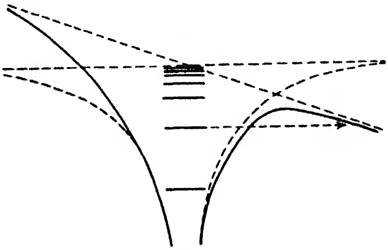


FIG. 105. Potential curve $U(r)$ and energy levels for a hydrogen atom in a strong external electric field. Dotted curve = undisturbed atom, solid curve = potential distorted by external field causing ionization of the atom in all states above the maximum of the potential curve at the right.

In conclusion, we discuss the third possibility of limiting the mean life of stationary states, that of collisions with gas particles of all kinds, by which the energy is carried away in collisions of the second kind, or by which, quite generally, the emission process is disturbed. The result of the reduction of the mean life of the stationary states by collisions is the so-called collision broadening of spectral lines. The investigation of this phenomenon has developed, in the last 20 years, into an independent and important branch of spectroscopy. Because of the difficulty in defining collisions (as discussed on page 19) there is, as can be shown exactly, no sharp distinction between the direct collision broadening resulting in a reduction of the mean life of the excited states, and the perturbation or broadening of atomic states by the inhomogeneous electric fields of neighboring electrons and ions. By an interatomic Stark effect such interatomic fields produce a variation in the amount of the displacement and spacing of the terms and thus a displacement and separation of the spectral lines which, due to statistical effects, appear as line broadening. In favorable cases the density of the perturbing electrons can be computed by means of the above theory from measurements of the line widths. Since in thermal equilibrium, say in an arc discharge, the electron density depends essentially on the temperature, the measurement of line widths thus furnishes a method of determining very high temperatures, in the range above 6000°K. Generally the perturbability of different atoms varies considerably, and it is apparent that the H atom and the alkalies are specially perturbable because their one external electron is at a great distance from the nucleus or core, respectively. Consequently, the spectra of these atoms exhibit the greatest line broadening. The broadening can become so extensive with these atoms that the different lines flow together to form one continuous spectrum. Figure 106 is an example of a spectrum of a spark discharge in compressed hydrogen, in which the broadened Balmer lines merge into a single continuous spectrum as the gas pressure is increased.

The extreme case of such perturbation of an atom by its neighbors exists in liquids and solids, which will be discussed from the atomic viewpoint in Chap. 7. Disregarding certain exceptions, which will be explained theoretically later, we no longer have, in these cases, sharp spectral lines, but only more or less extended continuous spectra or spectral bands as a result of perturbation by the environment.

Literature

General Literature

Condon, E. U., and G. H. Shortley: The Theory of Atomic Spectra, Cambridge University Press, London, 1935.

Grotrian, W.: Graphische Darstellung der Spektren von Atomen und Ionen, 2 vols., Springer-Verlag, Berlin, 1928. (Reprinted by Edwards Brothers.)

Herzberg, G.: Atomic Spectra and Atomic Structure, Dover, New York, 1946.

Hund, F.: Linienspektren und periodisches System, Springer-Verlag, Berlin, 1927.

Kuhn, H.: Atomspektren, vol. 9/I, Hand- und Jahrbuch der Chemischen Physik, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1934.

Pauling, L., and S. A. Goudsmit: The Structure of Line Spectra, McGraw-Hill Book Company, Inc., New York, 1930.

Schäfer, Cl.: Quantentheorie, vol. III/2 of the Lehrbuch der Theoretischen Physik, W. de Gruyter, Leipzig, 1937.

Sommerfeld, A.: Atombau und Spektrallinien, 6th ed., vol. I, Vieweg, Brunswick, 1944.

White, H. E.: Introduction to Atomic Spectra, McGraw-Hill Book Company, Inc., New York, 1934.

Special Literature

SECTION 1:

Bomke, H.: Vakuumspektroskopie, J. A. Barth, Leipzig, 1937.

Gordy, W.: Microwave spectroscopy, *Rev. Mod. Phys.*, **20,** 668 (1948).

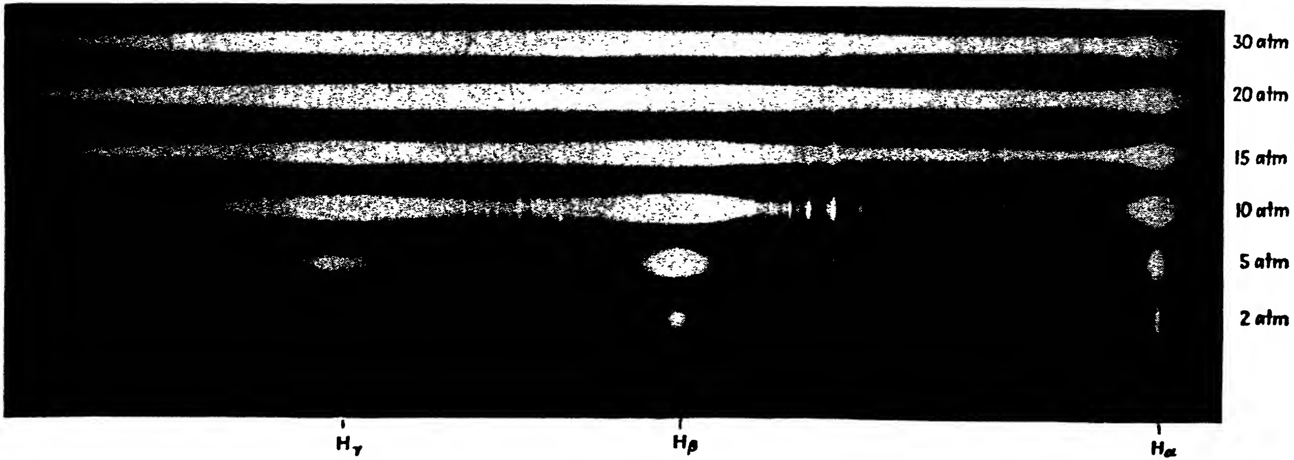


FIG. 106 Transition of the Balmer lines $H\alpha$, $H\beta$, and $H\gamma$ into a uniform emission continuum as a result of the effect of strong interatomic electric fields. Spectrum of a spark in hydrogen of 2 to 30 atmospheres pressure.

Harrison, G.: Wavelength Tables with Intensities in Arc, Spark or Discharge Tubes of More than 100,000 Spectrum Lines, John Wiley & Sons, Inc., New York, 1939.

Harrison, G. R., R. C. Lord, and J. R. Loofbourow: Practical Spectroscopy, Prentice-Hall, Inc., New York, 1948.

Kayser, H., and H. Konen: Handbuch der Spektroskopie, 7 vols., S. Hirzel, Leipzig, 1900–1934.

Kellogg, J. B. H., and S. Millman: Radiofrequency spectra of atoms and molecules, *Rev. Mod. Phys.* **18,** 323 (1946).

Siegbahn, M.: Spektroskopie der Röntgenstrahlen, 2d ed., Springer-Verlag, Berlin, 1931.

SECTION 2:

Bacher, R., and S. Goudsmit: Atomic Energy States, McGraw-Hill Book Company, Inc., New York, 1932. (In the process of being supplemented by a publication from the National Bureau of Standards: "Atomic Energy Levels.")

Paschen, F., and R. Goetze: Seriengesetze der Linienspektren, Springer-Verlag, Berlin, 1922.

SECTION 4:

Franck, J., and P. Jordan: Anregung von Quantensprüngen durch Stösse, Springer-Verlag, Berlin, 1926.

Penning, F. M., and W. de Groot: Anregung von Quantensprüngen durch Stösse, vol. 23/I, 2d ed., Handbuch der Physik, Springer-Verlag, Berlin, 1933.

SECTION 6:

Finkelnburg, W.: Kontinuierliche Spektren, Springer-Verlag, Berlin, 1938. (Reprinted by Edwards Brothers.)

Mitchell, A. C. G., and M. W. Zemansky: Resonance Radiation and Excited Atoms, The Macmillan Company, New York, 1934.

Mott, N. F., and H. S. W. Massey: The Theory of Atomic Collisions, Oxford University Press, New York, 1933.

SECTION 8:

Siegbahn, M.: Spektroskopie der Röntgenstrahlen, 2d ed., Springer-Verlag, Berlin, 1931.

SECTION 12:

Moore, C. E.: A Multiplet Table of Astrophysical Interest, Princeton University Press, Princeton, N.J., 1945.

SECTION 13:

Stoner, E. C.: Magnetism and Matter, Methuen & Co., Ltd., London, 1934.

Van Vleck, J. H.: The Theory of Electric and Magnetic Susceptibilities, Oxford University Press, New York, 1932.

SECTION 14:

Back, E., and A. Landé: Zeemanneffekt und Multiplettstruktur, Springer-Verlag, Berlin, 1925.

SECTION 17:

Kopfermann, H.: Kernmomente, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1941.

SECTION 18:

Finkelnburg, W.: Kontinuierliche Spektren, Springer-Verlag, Berlin, 1938.

CHAPTER 4

THE ATOMIC THEORY ACCORDING TO QUANTUM MECHANICS

4-1. The Transition from Bohr's Half-classical Theory to the Quantum-mechanical Theory of Atoms

The theoretical treatment of atomic physics as it was presented in the preceding pages was based on the introduction of simple atomic models. The behavior of the models was governed by the known laws of classical physics. These laws were then used to obtain the results described in the last chapter. These classical theoretical results were "forced" into agreement with experimental results by introducing arbitrary and basically incomprehensible quantum conditions.

The amazing success of the "naïve" theory presented in the last chapter is indisputable and shows that, in spite of its simplicity, the Bohr theory contains a great amount of truth. On the other hand, not only is the arbitrary imposition of quantum conditions a very unsatisfactory feature of the Bohr theory, but the theory fails completely in a number of outstanding examples, such as the model of the helium atom; the derivation of the Landé factor for the anomalous Zeeman effect (page 161); the explanation of the "Pendelbahnen" with $l = 0$ (in which the electrons with orbital momentum zero should oscillate through the nucleus, which obviously is impossible!); and in the case of the quantitative theory of molecular spectra and chemical binding, which will be considered later (page 409). Since about 1924 physicists were very much aware that all these difficulties and inadequacies could be removed only by developing a new atomic theory which must essentially agree with the results of the naïve theory, and at the same time be a definite departure from any classical basis. The development of the new atomic theory was carried out simultaneously and independently in 1926 by Schrödinger, who based his work on that of de Broglie, and by Heisenberg, who used a more fundamental approach. Both theories, wave mechanics and quantum mechanics, were, as was soon realized, basically identical in spite of different initial assumptions and mathematical treatment. They constitute a definite rupture with the earlier, supposedly self-evident conception, that microphysical phenomena must be described in the same manner and with the same laws used in classical physics to describe the behavior of systems which are large compared with atomic dimensions. We shall consider later how this new conception of microphysics modified all scientific thinking and, consequently, natural philosophy.

In the beginning the Heisenberg-Schrödinger quantum mechanics had two deficiencies. It did not take into account the electron spin or the theory of relativity. Soon, however, the first defect was eliminated by Heisenberg, Jordan, and Pauli; later Dirac developed the theory to encompass the more complete quantum mechanics and quantum electrodynamics. Dirac's theory is beyond the scope of this book so we shall restrict ourselves to the simpler form, i.e., we shall disregard the theory of relativity and omit quantum electrodynamics, but we shall discuss unrelativistic wave mechanics and then supplement it by introducing the electron spin.

The rigorous development of quantum mechanics, as we shall call the new theory, led to a more complicated but conceptually consistent atomic theory, which, to "normal" thinking, appears to be rather strange. But the new theory not only completely overcame the objections to the old atomic theory but also led to many new results which were confirmed by experiment. We shall deal with these results in all the following chapters. Thus we have good reason to believe that this theory provides an accurate theoretical description of atomic processes and so may be considered the correct atomic theory. Only with the most extreme phenomena of nuclear physics, e.g., very high energy collision processes which we will consider later (page 322), do we find that even quantum mechanics is not general enough. So we may expect further research to lead to an even more general theory which includes quantum mechanics as a special case.

In the following discussion, the fundamental reasoning and basic equations of wave mechanics will be presented in a form especially suited for calculating atomic phenomena. In this respect we restrict ourselves in the mathematical development to considering only what is necessary to understand the conclusions. We shall emphasize, however, as clearly as we are able to, the basically new features of the theory and its implications with regard to the wider field of atomic physics including nuclear physics, molecular physics, and physics of the solid state.

Since the new atomic theory is based on the wave-particle dualism of light and matter, we shall discuss this next.

4-2. The Wave-Particle Dualism of Light and Matter

From the time of Huygens up to the establishment of the Bohr atomic theory there was not the slightest doubt that light was an electromagnetic wave phenomenon and that, on the other hand, matter in the form of molecules, atoms, and electrons had to be regarded as particles which, in the first approximation, could be treated as point masses. The development of the quantum-mechanical atomic theory began when it was realized that the conception of waves and particles as being unrelated entities, once considered self-evident, was too restricted. Particularly, the manner in which experiments are performed with radiation and matter determines whether they display wave or particle characteristics. Wave and particles are but two manifestations of the same physical reality.

The wave nature of light seemed so definitely proved by numerous interference and diffraction experiments that the demonstration of the diffraction of X-rays by von Laue was regarded as a definite proof of the wave nature of X-rays and, consequently, of the identity of X-rays and light. The first light phenomenon observed which could not be explained by the wave theory was the photoelectric effect, the release of electrons from a metal surface when irradiated by light (page 31). The experimental result, that the velocity of the electrons escaping from the metal depends only on the color (wavelength or frequency) of the incident light and not on its intensity, could not be understood from the standpoint of the wave theory of light, according to which the energy of any radiation is proportional to the square of the amplitude of the waves, but has nothing to do with their frequency. The correct relation between the mass and velocity of the escaping electrons, on the one hand, and the frequency ν of the incident light on the other, was derived by Einstein from experimental results. He found

$$\frac{m}{2}v^2 = h\nu - \phi \tag{4-1}$$

where ϕ, the work function, is a characteristic constant for each metal (page 445), and, to the surprise of all investigators, the constant h proved to be identical with the elementary quantum h introduced by Planck in his theory of heat radiation, as mentioned on page 64. Einstein concluded that light actually consisted of individual quanta of energy $h\nu$, now called photons. Accordingly, photons must be emitted from the source into all space and they must also be capable of being absorbed as quanta by atoms. For the absorption of light had also become a scarcely understandable phenomenon from the point of view of the wave theory. One has only to ask how it is possible that a spherical wave expanding into space can be absorbed in its entirety by *one* atom. The photon hypothesis automatically removes this obstacle.

The decisive experiment with respect to the wave or particle nature of electromagnetic radiation was the discovery of a famous effect in 1922 by A. H. Compton. Shortwave X- or γ-rays of definite wavelength are made to impinge on a suitable solid body such as graphite, and the laterally scattered radiation is observed. Spectral analysis reveals that in addition to a line corresponding to the scattered primary wavelength there is another line displaced in the direction of longer wavelengths. This wavelength difference between primary line and scattered Compton line increases with increasing angle of observation (measured from the direction of the incident ray). From the particle concept one can regard the process as a collision of a photon with an electron of an atomic shell, whereby the photon transfers momentum and energy to the electron. By the impact of the photon, the electron is kicked out of its atom and deflected into a direction in accordance with the laws of elastic impact (Fig. 107). For the theoretical treatment we need the energy E and momentum p of the photon which were discussed in Chap. 2,

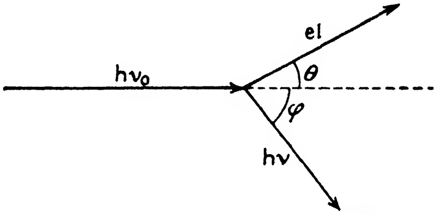


FIG. 107. Compton scattering of a photon $h\nu_0$ through the angle φ in a collision with an electron (Compton effect).

$$E = mc^2 = h\nu \tag{2-51}$$

$$p = mc = \frac{h\nu}{c} \tag{2-54}$$

Let ν_0 and ν be the frequency of the photon before and after its collision with an electron of mass m_e, respectively, which may, after the impact, have a velocity v. Then we obtain, for the energy and the components of the momentum parallel and perpendicular to the incident ray, the following equations:

$$h\nu_0 = h\nu + \frac{m_e}{2}v^2 \tag{4-2}$$

$$\frac{h\nu_0}{c} = \frac{h\nu}{c}\cos\varphi + m_e v \cos\vartheta \tag{4-3}$$

$$0 = \frac{h\nu}{c}\sin\varphi + m_e v \sin\vartheta \tag{4-4}$$

By eliminating the angle ϑ and the electron velocity v, which are of no interest, from these three equations, we get

$$h\nu_0 = h\nu + \frac{h^2}{2m_e c^2}(\nu_0^2 + \nu^2 - 2\nu_0\nu\cos\varphi) \tag{4-5}$$

By solving for $\nu_0 - \nu = \Delta\nu$ and substituting ν for ν_0 in the expression in parentheses, since $\Delta\nu$ is small compared to ν and ν_0, we arrive at the following equation

$$\Delta\nu = \frac{2h\nu^2}{m_e c^2} \sin^2 \frac{\varphi}{2} \tag{4-6}$$

It gives the frequency difference between the incident γ-ray and that scattered by the Compton effect as a function of the angle between the incident and the scattered photon (Fig. 107). By converting the $\Delta\nu$ into wavelength differences, we finally have

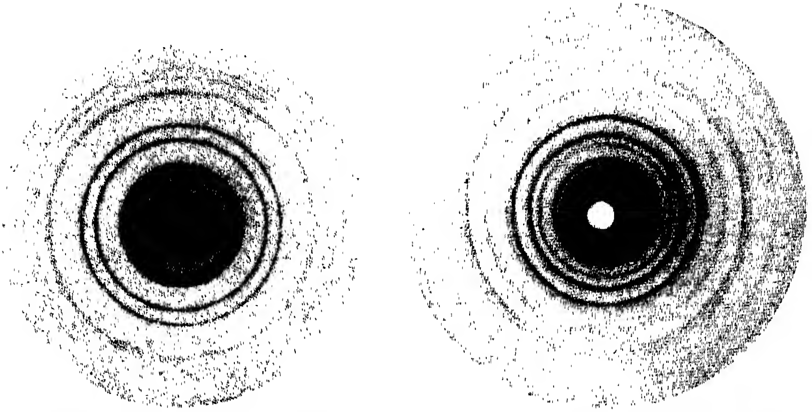
$$\Delta\lambda = \frac{2h}{m_e c} \sin^2 \frac{\varphi}{2}. \tag{4-7}$$

This last equation is in complete agreement with experimental results. Sometimes the quantity $h/m_e c = \lambda_e$, which has the dimension of length, is called the "Compton wavelength" of the electron. The wavelength difference between the scattered Compton line and the primary line is, from (4-7), independent of the wavelength of the primary line and depends only on the angle of deflection φ. Although this derivation did not take into consideration the relativistic mass dependence, it agrees exactly with the results of the more rigorous relativistic derivation.

Compton scattering thus can be shown to be in quantitative agreement with the experimental results if the photon under consideration is assumed to have particle properties so that, like a billiard ball, it gives up energy and momentum to the electron with which it collides. The photon then is deflected, according to the laws of elastic impact, through an angle φ from its original direction, and its loss of energy can be determined as a function of φ by measuring the wavelength shift, $\Delta\lambda$, with an X-ray spectrograph.

We thus know about light phenomena, such as interference and diffraction, which can be explained only by the wave theory, and of other phenomena, such as the photoelectric effect, the absorption of light, and the Compton effect, which can be explained easily only by the particle concept. *The earlier assumption, that the wave and particle concepts were mutually exclusive, has yielded today to the conviction that light can be considered as an extended field of waves* or *as point-like particles depending upon the manner in which experiments with light (radiation) are performed.* Later on we shall thoroughly investigate how this conclusion, forced on us by experimental results, leads to a theory that is consistent and which is adequate to describe nearly all atomic phenomena. This theory, however, is not pictorial like the Bohr theory, because the concepts of extended fields of waves and point particles seem to be irreconcilable contradictions, at least from our present point of view.

The converse of our statements about radiation holds for matter. The phenomena of the kinetic theory of gases seem to provide such clear evidence for the existence of corpuscular molecules and atoms, and numerous cathode-ray experiments have so plainly shown that the electron with its mass and velocity appears as a definitely corpuscular particle, that it seemed almost absurd, before 1924, to question the particle nature of matter.



(a) Diffraction of 36,000-volt electrons ($\lambda = 0.06$ A).

(b) Diffraction of X-rays (copper K_α radiation $\lambda = 1.54$ A).

FIG. 108. Comparison of electron diffraction and X-ray diffraction by a silver foil. (*After Mark and Wierl.*)

However, following the theoretical deliberations of de Broglie, which will be reviewed shortly, Davisson and Germer as well as G. P. Thomson obtained diffraction patterns by directing beams of electrons onto crystals. These patterns, except for small explainable deviations due to a difference of the refractive index, are completely analogous to the X-ray diffraction patterns of von Laue or Debye-Scherrer. Figure 108 shows an example of electron diffraction by a thin metal foil, and, for comparison, a picture of an X-ray diffraction of the same material. These experiments, interpreted in the customary way, indicate a wave nature of the electron. The technique of electron diffraction has, in the meantime, become common knowledge to physicists and chemists, who use electron as well as X-ray diffraction to a large extent as a means of determining molecular and crystal structures (pages 349 and 426). From electron diffraction measurements on crystal lattices of known lattice constants, essentially as in X-ray diffraction, the wavelength of the "material waves" corresponding to a beam of electrons can be determined. These measurements gave the astounding result that this wavelength of electrons or any other particles depends on the mass and velocity of the particles in question, and that the value of this so-called de Broglie wavelength of material particles is

$$\lambda = \frac{h}{mv} \tag{4-8}$$

For electrons having a velocity v, due to a potential difference,

$$U = \frac{mv^2}{2e} \qquad v = \sqrt{\frac{2eU}{m}} \tag{4-9}$$

we get, by substituting the values for m_e and e,

$$\lambda = \frac{12.3}{\sqrt{U}}\,\mathrm{A} \tag{4-10}$$

where U is measured in volts. For example, an electron which has fallen through a potential difference of 10,000 volts has a wavelength of 0.12 A, corresponding to that of hard X-rays.

Diffraction experiments illustrating the wave nature of matter can be performed successfully not only with electrons but also with atomic and molecular beams. However, since the mass appears in the denominator of Eq. (4-8), the corresponding wavelength of these heavier particles for the same velocity is smaller by a factor of about 10^{-4}. Using the comparatively small velocity of atoms and molecules due to their random thermal motion, Stern carried out diffraction experiments also with these heavier particles and was able to verify Eq. (4-8). *All these diffraction experiments provide definite evidence for the extremely important fact that in suitable experiments particle beams of electrons, atoms, and molecules exhibit a wave-like character.* That this wave character of matter remained undiscovered so long, and that it plays no part in macrophysics, follows from Eq. (4-8) for the de Broglie wavelength. Macrophysics deals throughout with such large masses that moving bodies have, according to Eq. (4-8), associated wavelengths by far too short to be detectable. *Macrophysics and its laws, therefore, are not changed by the new results or the theoretical conclusions which will be drawn from them in the next sections, in spite of the fundamental change in our conception of physics.* The apparently so fundamental difference between light waves on the one hand and material particles on the other has largely disappeared. *Light (radiation) as well as matter exhibit wave or particle characteristics depending on the kind of experiments which are performed. For both of them there is no either/or, but characteristic of both is the duality of phenomena as described above. It is the foundation of the new atomic physics.*

4-3. Heisenberg's Uncertainty Principle

Before considering the significance of the wave nature of the electron in atomic physics and familiarizing ourselves with the characteristic features of the wave-mechanical formalism, we shall discuss a fundamentally important consequence of the wave-particle dualism which we have described.

We ask the question: how can a physical system, for whose complete physical description wave and particle properties are necessary, be described unambiguously and exactly with respect to its behavior in space and time? In investigating this problem, we shall find a very peculiar and significant uncertainty, discovered by Heisenberg in 1926, which applies to simultaneous statements on position and momentum, or energy and time, of the system.

Let us discuss first the case of light and think of an infinitely long wave train whose wavelength can be measured with a spectrograph to a fundamentally arbitrary accuracy. From the wavelength, we obtain, by means of the well-known relation

$$\nu = \frac{c}{\lambda} \tag{4-11}$$

with equal accuracy, the frequency ν and, by means of the fundamental equation of the quantum theory (Bohr's frequency condition), the energy

$$E = h\nu = \frac{hc}{\lambda} \tag{4-12}$$

as well as, finally, the momentum

$$p = mc = \frac{h\nu}{c} = \frac{h}{\lambda} \tag{4-13}$$

of the corresponding photons in the particle picture. Thus, by means of an infinitely long monochromatic wave train, the momentum of the associated photons is exactly determined; it is to be regarded from the standpoint of our derivation as a *wave property*. However, we note that no statement is possible in this case about the position of a particular photon, since the wave train (or by transition to three coordinates, the three-dimensional wave field) extends over the whole space and no coordination is possible between the corresponding photons and their particular positions in space.

Conversely, to determine the position of a photon we must, using the wave presentation, go from an infinitely extended wave train (Fig. 109*a*) to an increasingly shorter wave train, which in the limiting case is reduced to a single maximum (Fig. 109*b*). Evidently then the corresponding photon must be sought for at the location of this maximum, so that its position then would be known accurately. However, now the attempt to determine *simultaneously* the momentum of the photon from Eq. (4-13) by measuring the wavelength fails, since the Fourier analysis, by means of the spectrometer which is necessary for this measurement, reveals a continuous spectrum with all wavelengths between 0 and ∞ (though of course with varying intensities). The wavelength and the momentum of the photon associated with the wave pulse thus cannot be determined in this case in which the position of the photon is known accurately.

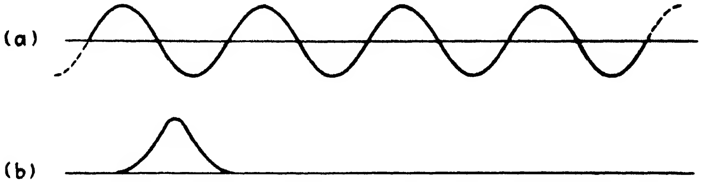


Fig. 109. (*a*) Infinitely extended wave train and (*b*) single wave pulse.

We now show quantitatively by a discussion of a simple diffraction experiment the important fact of the impossibility of simultaneously determining position and momentum. Monochromatic light of wavelength λ may be diffracted through an angle φ by a slit of width b (Fig. 110). From optics it is known that for a given λ, b, and φ, the first diffraction maximum is given by the relation

$$\lambda \approx b \sin \varphi \qquad (4\text{-}14)$$

Thus the smaller the slit width b, the larger is the diffraction angle φ. If we go from the wave picture of diffraction over to the particle picture, the slit width b determines the position (coordinate q) of a photon which passes the slit, up to an uncertainty Δq of the order of the slit width,

$$\Delta q = b \qquad (4\text{-}15)$$

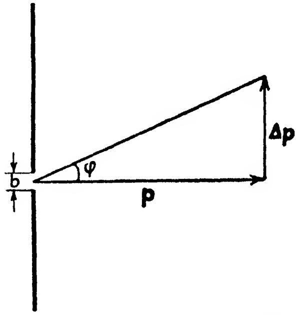


Fig. 110. Diffraction by a slit as an example of the uncertainty principle. b = slit width, φ = diffraction angle, $\mathbf{p}$ = initial momentum of the photon, $\Delta\mathbf{p}$ = change of momentum of the photon by the diffraction.

The diffraction of a wave through an angle φ corresponds in the particle picture to a change of the particle momentum (which is a vector) by the amount Δp. If we assume that

$$\Delta p \ll p \qquad (4\text{-}16)$$

then φ is small and

$$\Delta p = p \tan \varphi \approx p \sin \varphi \tag{4-17}$$

By substituting $\sin \varphi$ from (4-14) and using (4-13) we have

$$\Delta p \approx \frac{p\lambda}{b} \approx \frac{h}{b} \tag{4-18}$$

and finally from (4-15)

$$\Delta p \Delta q \approx h \tag{4-19}$$

Consequently, in agreement with the result of the above simple treatment, position and momentum (velocity) of a particle cannot be simultaneously determined with any accuracy. *The product of the error in measuring the position and that of measuring the momentum is at least of the order of magnitude of Planck's constant h.* An exact calculation by Heisenberg gave $h/4\pi$ as the minimum value. In this connection we speak of "at least" because in addition to this fundamental uncertainty we always have an uncertainty caused by the inevitable errors of measurement. This is the famous *Heisenberg uncertainty principle,* which generally applies to so-called canonically conjugated variables such as position and momentum. Quite generally, canonically conjugated variables are such related quantities whose product has the dimensions of an action (Energy × time = gm cm^2 sec^{-1}).

Thus energy and time are also canonically conjugated variables. Because these have a special significance in atomic physics, the validity of the uncertainty principle may be derived briefly for them also. In our first example (page 199) with wave trains of different lengths, the accuracy with which the time can be determined, for instance the transit of a photon through a fixed plane, is the greater and, accordingly, the time error the smaller, the shorter the length of the wave train. However, the shorter the length of the wave train, the greater, according to the Fourier analysis, is the frequency range of the waves forming the wave train. Consequently

$$\Delta t \approx \frac{1}{\Delta \nu} \tag{4-20}$$

or from the basic equation of quantum theory

$$\Delta E = h\Delta \nu \tag{4-21}$$

$$\Delta t \Delta E \cong h \tag{4-22}$$

The uncertainty principle thus holds in this case also. This means that the more exact the time of an atomic process is determined, or the shorter the duration of the process, the less exact is our possible knowledge of the energy change of the atomic system under consideration.

An important example for this form of the uncertainty principle is the normal width of spectral lines which we have already considered (page 185). The smaller the mean life $\tau = \Delta t$ of an excited atomic state is, the less exactly its energy is determined, the greater is the width ΔE of this state and consequently the width of the spectral line resulting from a transition from that state.

In the above derivations we have dealt exclusively with the case of light. However, the results are universally valid for all measurements of canonically conjugated variables of any physical system. We treat as a further example the famous imaginary attempt of measuring simultaneously position and velocity of an electron with an imaginary supermicroscope. Its objective lens may have an aperture angle u (Fig. 111). Then the inaccuracy Δq resulting from the determination of the position of the electron by illuminating it with light of wavelength λ is given, according to the laws of wave optics, by the resolving power of the objective and is

$$\Delta q = \frac{\lambda}{\sin u} \qquad (4\text{-}23)$$

Now the position of an electron can be measured with a microscope only if, at the very least, one photon is deflected by the electron from its primary direction (Fig. 111) into the microscope. However, by this deflection (scattering), because of the Compton effect (page 194), the electron receives a recoil of the order of magnitude of $h\nu/c$. This change of the momentum Δp is not exactly determinable, because all that is known about the scattered photon is that it is scattered into the solid angle u (Fig. 111). So the uncertainty of the momentum is

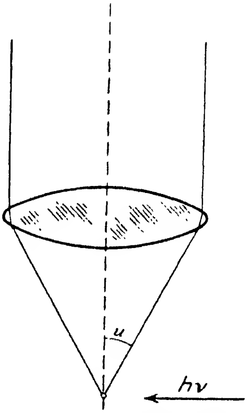


FIG. 111. Imaginary experiment concerning the uncertainty principle: microscopic observation of an electron. Arrow $h\nu$ indicates the direction of the incident light; u = half angle of the microscope lens. (*After Heisenberg.*)

$$\Delta p = \frac{h\nu}{c} \sin u \qquad (4\text{-}24)$$

and from (4-23) and (4-24) follows finally

$$\Delta p \Delta q = h \qquad (4\text{-}25)$$

Because of the fundamental significance of the uncertainty principle, we shall discuss it finally from an entirely different point of view.

Since Planck's discovery (page 64) we know that the physical quantity of action is not infinitely divisible in nature, but occurs in smallest quantities which are called the elementary quanta of action h. Thus, if we try to measure simultaneously in a physical system two canonically conjugated quantities such as energy and time, or position and momentum, whose product has the dimension of an action, the product of these two quantities cannot be determined more accurately than within *one* action quantum h, simply because there *is* no fraction of h. *From this deliberation follows at once the uncertainty relation* (4-22) *or* (4-25) *as a logical consequence of the existence of the elementary quantum of action.*

The last consideration shows most clearly that *the Heisenberg uncertainty principle is not the result of any inadequacy of our ability of measuring, but is a logical consequence of Planck's quantization of action, and so establishes a fundamental limit of accuracy for the measurement of any canonically conjugated quantities of a system.*

4-4. De Broglie's Particle Waves and Their Significance for the Bohr Atomic Theory

The idea of the wave character of material particles, conceived by de Broglie in 1924, was based on the wave particle dualism of light which had led to the reasonable question whether a similar dualism might also exist for matter. In the case of light, the relation of the energy E and the momentum p of a photon to the frequency ν and the wavelength λ of the corresponding wave is given by the following equations which we have already used (page 62):

$$E = h\nu = \frac{hc}{\lambda} \tag{2-51}$$

$$p = mc = \frac{h\nu}{c} = \frac{h}{\lambda} \tag{2-54}$$

De Broglie made the attempt to apply these formulas to material particles of mass m traveling with velocity v. By substituting the particle mass and velocity into Eq. (2-54) he immediately obtained the wavelength associated with such a particle,

$$\lambda = \frac{h}{p} = \frac{h}{mv} \tag{4-26}$$

We have already learned on page 196 that this result was confirmed somewhat later in the most startling way by the electron diffraction experiments of Davisson, Germer, Thomson, and others.

The velocity of propagation of the material waves u, which has to be distinguished from the particle velocity v, we get from (4-26) and (2-51)

$$u = \lambda\nu = \frac{E}{mv} = \frac{mc^2}{mv} = \frac{c^2}{v} \tag{4-27}$$

The phase velocity u, as the velocity of propagation of the material wave is called to distinguish it from the particle velocity v, is thus inversely proportional to the particle velocity v. This result (4-27) is most surprising. As the particle velocity v is always smaller than the velocity of light c, according to (4-27) the phase velocity of the de Broglie waves is always *larger* than the velocity of light. This result is not in contradiction to the relativity theory, as one might think offhand, because the theory requires only that there be no velocity useful for the transmission of signals which is greater than c. And the phase velocity of waves, because of the complete identity of their wave peaks, cannot be used for signal transmission.

It is of interest that the material wave concept includes as a limiting case light waves. Namely, by applying the theory and (4-27) to particles which are traveling with the velocity $v = c$, we get the result that the particles must have a rest mass equal to zero and thus are identical with photons. To prove this we proceed from the momentum and energy formulas (2-51) and (2-54), which, after taking into account the relativistic mass dependency, become

$$p = \frac{m_0 v}{\sqrt{1 - (v^2/c^2)}} \tag{4-28}$$

$$E = h\nu = \frac{m_0 c^2}{\sqrt{1 - (v^2/c^2)}} \tag{4-29}$$

By eliminating v and solving for the rest mass, m_0, we find

$$m_0 = \frac{h}{c}\sqrt{\left(\frac{\nu}{c}\right)^2 - \frac{1}{\lambda^2}} \tag{4-30}$$

For particles which are propagated with the velocity of light the wavelength is $\lambda = c/\nu$ and substitution in (4-30) gives the result

$$m_0 = 0 \tag{4-31}$$

The fact that the formalism indicated above comprises the wave properties of material particles as well as photons in agreement with experience, is a strong point in favor of the new theory.

Now, how does one describe the wave-particle duality from the viewpoint of wave mechanics, as de Broglie called the new theory? A wave train of wavelength λ given by (4-26) corresponds to a beam of particles of the same velocity v. The direction of propagation of the material waves corresponds to the beam direction in the particle picture. What was said in the section on the uncertainty principle holds for the relation of the wave characteristics to the particle properties. An extended wave field (or, in one dimension, a wave train) permits the exact determination of the wavelength of the particle wave and, consequently, the momentum (and the velocity) of the corresponding particles. A specification of the position of the particle is in this case impossible. A locally determined particle, however, corresponds in the wave picture to a *wave packet* formed by superposition of *many* waves of somewhat different frequency. According to the general wave theory such a wave packet is propagated with the so-called group velocity which is different from the phase velocity u and can be proved to be identical with the particle velocity v. The more accurately it is possible to determine the position of the wave packet, the more numerous are the component frequencies which form the packet and are revealed by a Fourier analysis, the less exactly, on the other hand, can a wavelength, and the momentum and velocity of the associated particle, be determined (page 201).

Let us try to understand the full significance of this correlation by considering it from a somewhat different point of view. The same physical phenomenon, e.g., a cathode ray consisting of electrons of given velocity v, can be described in two apparently irreconcilable ways: as a number of dimensionless point masses moving with the velocity v, and as an infinitely extended field of material waves of wavelength $\lambda = h/mv$ with no indication of single masses at definite positions. However, our experimental knowledge tells us that in all practical cases this last conception is far too extreme. The location of an electron is, evidently, at the very least confined to the discharge tube or the cathode ray. This confinement with respect to position, according to the uncertainty principle, implies an uncertainty Δp for the electron momentum which is inversely proportional to the positional confinement. This indefiniteness of p means, according to the de Broglie relation (4-26), a scattering of the wavelengths of the electron waves about a most probable mean value. By the superposition of these somewhat varying wavelengths a *wave group* is formed, a *wave packet*, which is propagated with a group velocity v equal to the particle velocity. Such a wave packet has an amplitude differing essentially from zero only in a very limited region of space, in contrast to the infinitely extended monochromatic waves. Thus we obtain within the wave picture a much better representation of a beam of electrons than by the unsatisfactory infinitely extended de Broglie wave (4-26) with exactly fixed wavelength, at which we would have arrived by a superficial formal derivation. The wave packet concept stands, as we see, to a certain degree between the two extremes of the dimensionless particle and the infinitely extended wave.

The described considerations of de Broglie of the wave characteristics of matter lead to an important new interpretation of the basic postulates of the Bohr theory (see page 84) with respect to the discrete quantum orbits of the atomic electrons. For if the electron is considered as a wave, a stationary orbit such as in Fig. 112 is distinguished from any arbitrary orbit by the fact that its circumference must be an integral multiple of the wavelength of the electron. Otherwise the electron wave would destroy itself in time by interference, and no stationary state of the atom would be possible. Actually, de Broglie was able to show that the arbitrarily introduced quantum condition for the principal quantum number n (page 85) is identical with the requirement that the circumference of the orbit be an integral multiple of the de Broglie wavelength of the revolving electron. For if we write the Bohr quantum condition according to Eq. (3-6) for the special case of a circular orbit as

$$2\pi rmv = nh \qquad n = 1, 2, 3, \ldots \tag{3-6}$$

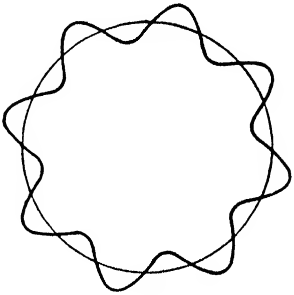


FIG. 112. Schematic representation of a (one-dimensional) standing electron wave in a Bohr orbit.

and substitute for the momentum mv of the electron according to Eq. (4-26) the de Broglie wavelength

$$mv = \frac{h}{\lambda} \tag{4-32}$$

we obtain

$$2\pi r = n\lambda \qquad n = 1, 2, 3, \ldots \tag{4-33}$$

which was to be proved. The proof can be extended, with somewhat more mathematical expenditure, to the general case of elliptical orbits. The Bohr quantum condition (3-6), at first entirely arbitrary, thus is reduced to a physically significant condition of stationarity whereby the standing electron wave takes the place of the electron traveling around the orbit. It may be pointed out here already that this concept of an orbital electron as a standing wave also removes the much emphasized difficulty (page 84) that the electron travels in its orbit without radiating, contrary to the requirement of classical electrodynamics. If the electron actually does not, in the customary sense, rotate as a point charge and thus does not form an oscillating dipole, but as a standing wave forms a stationary time-independent system, the classical requirement of dipole radiation does not apply. The failure of the stationary orbits to radiate, and especially the stability of the ground orbits, thus is explained without contradicting the usual concepts of electrodynamics. We will consider this in more detail shortly (page 214).

This presentation so far does not account for what takes place in the space outside the actual Bohr orbits. The answer to this question together with all other detail features of wave mechanics follows from the more vigorous development of the theory.

4-5. The Basic Equations of Wave Mechanics. Eigenvalues and Eigenfunctions. Matrix Mechanics and Its Relation to Wave Mechanics

We proceed now to this more rigorous discussion of the wave picture of matter and thus associate with a particle a wave phenomenon Ψ. The exact physical meaning of this quantity Ψ which varies periodically in time and space and which expands with the phase velocity u will be discussed later. This Ψ-wave, which Schrödinger introduced in further developing the ideas of de Broglie, must satisfy the general wave equation

$$\nabla^2\Psi = \frac{1}{u^2}\frac{\partial^2\Psi}{\partial t^2} \tag{4-34}$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{4-35}$$

is the so-called Laplace operator. For the velocity of propagation u of this Ψ-wave we use, according to Eq. (4-27)

$$u = \frac{h\nu}{mv} \tag{4-36}$$

The validity of this equation follows from the agreement of wave mechanics with the empirical results of atomic physics. If we substitute for the particle momentum mv an expression containing the total energy E and the potential energy U of the system according to the relation

$$\frac{m}{2}v^2 = E - U \tag{4-37}$$

we obtain for the phase velocity

$$u = \frac{h\nu}{\sqrt{2m(E - U)}} \tag{4-38}$$

and by substituting (4-38) in the general wave equation (4-34) the equation for the Ψ-wave

$$\nabla^2\Psi = \frac{2m(E - U)}{h^2\nu^2}\frac{\partial^2\Psi}{\partial t^2} \tag{4-39}$$

We now separate the quantity Ψ into two factors, one depending only on space coordinates and the other oscillating with a frequency ν, depending only on the time,

$$\Psi = \psi e^{-2\pi i\nu t} \tag{4-40}$$

in which ψ does not vary with time. By differentiating (4-40) twice with respect to the time, we obtain

$$\frac{\partial^2\Psi}{\partial t^2} = -4\pi^2\nu^2\Psi \tag{4-41}$$

and by substituting (4-41) in (4-39) under consideration of (4-40), we obtain *the famous Schrödinger equation*

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(E - U)\psi = 0 \tag{4-42}$$

which describes the space dependency of Ψ. Because of the Laplacian operator ∇^2, this is a homogeneous partial differential equation of the second order, which determines the behavior of the standing-wave phenomenon corresponding to a material system, if the potential energy of any special system is substituted for U. We will become acquainted with this procedure in detail in the next section.

Differential equations of the type of Eq. (4-42) are thoroughly familiar to physicists and mathematicians. They occur, for example, in the treatment of the elastic standing waves of the violin string or of the stretched membrane (drumhead). *It is characteristic for all of these cases that not* all *arbitrary vibrations* [*corresponding to arbitrary solutions* ψ *of Eq.* (4-42)] *can exist, but only certain definite wave forms and frequencies are possible. This is caused by the given boundary conditions, i.e., the fact that both ends of the string are held fast, or that the drumhead is fastened at its circumference.* This is easy to see in the one-dimensional case of the string. Since both ends of the string are fastened and thus are unable to move, they must be nodes of vibration and a stationary vibration is possible only if the length L of the string is an integral multiple of a half wavelength.

$$L = n\frac{\lambda}{2} \qquad n = 1, 2, 3, \ldots \tag{4-43}$$

We recognize the formal similarity of this condition for the stationary vibration of the vibrating string with the quantum condition (4-33) already considered for the electron oscillating in the Bohr orbit (page 205). The case of the vibrating membrane is more complicated, but fundamentally the same. Here also the circular fastening of the membrane, as on a drum, acts as a boundary condition that in general permits only discrete wave forms and frequencies. *Mathematically it turns out that the differential equation under consideration of the boundary condition (i.e.,* $\psi = 0$ *at the fixed circumference of the membrane) does not have solutions corresponding to stationary wave forms for all arbitrary frequencies, but only for certain "eigenvalues" or "eigenfrequencies" of the vibrating membrane. The wave forms associated with these eigenfrequencies are called the eigenfunctions of the vibrating membrane.*

From these macromechanical examples we understand why vibration problems which are characterized by certain boundary conditions are called boundary value problems or eigenvalue problems. These mathematical characteristics, which we also find in most applications of the Schrödinger equation, are common to all vibration problems and do not present, as the beginner often believes, a mysterious peculiarity of wave mechanics of atomic systems.

The parallel delineated above between the eigenvalues of the vibrating string and the quantum condition of the Bohr theory has also played a great role historically in the development of wave mechanics by Schrödinger, whose first paper bore the now understandable title "Quantization as an Eigenvalue Problem."

The Schrödinger equation (4-42) is, in general, also an eigenvalue equation, since we have for all stationary states of an atomic system, at least, the boundary condition, that the system and with it the corresponding wave phenomenon ψ is limited in space. This means that in infinity ψ must be zero. Thus in all stationary cases the Schrödinger equation has solutions only for certain eigenfrequencies ν. Because $E = h\nu$ these eigenfrequencies correspond to certain eigenvalues of the energy which are the stationary energy states of the system. The solutions of the Schrödinger equation, the eigenfunctions ψ which belong to these eigenvalues of the energy, characterize completely the behavior of the system in these energy states, e.g., that of an electron in the stationary states of an atom.

Let us consider, on the other hand, a beam of electrons (a cathode ray) moving in free space with a velocity v. The potential energy of this force-free system then is zero ($U = 0$), and the Schrödinger equation (4-42) (by neglecting the interaction of the electrons) has the simple form

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2} E\psi = 0 \qquad (4\text{-}44)$$

where E is equal to the kinetic energy of the cathode-ray electrons. This differential equation is not an eigenvalue problem since in this problem of free electrons no boundary condition limits the number of solutions. The theory of differential equations tells us that this equation has a solution for each arbitrary E value and that these solutions are sine functions. The treatment of the simple problem of freely moving electrons in the cathode ray by means of the Schrödinger equation consequently leads to the well known result that the material wave corresponding to freely moving electrons is a sine wave with a wavelength given by Eq. (4-26). As there evidently is no limitation for the electron velocity, there must be solutions for *all* E values.

For the case of stationary systems, which is of great interest in atomic physics, some further important statements can be made about the eigenfunction ψ. First of all we notice that in general the eigenfunctions are complex, but that a real quantity, the "norm" of the eigenfunction, is obtained by multiplying ψ by its complex conjugate ψ^*. We shall discuss the physical significance of this norm $\psi\psi^*$ in the next section (page 212). If for each eigenvalue of the system there is a single eigenfunction, i.e., wave form, then the system is called nondegenerate. If, on the other hand, there are n eigenfunctions belonging to the same energy eigenvalue, we speak of an $(n - 1)$-fold degeneracy.

We have mentioned already that the Schrödinger equation (4-42) is a homogeneous differential equation. Therefore, its solution is given except for a constant factor. In order to compare various eigenfunctions, they must be equated to each other. Then this factor is determined by the definition that the integral of the norm of the eigenfunction, taken over the whole space τ, shall be unity,

$$\int \psi\psi^* \, d\tau = \int |\psi|^2 \, d\tau = 1 \tag{4-45}$$

An eigenfunction determined in this manner is called normalized.

Furthermore, our eigenfunctions are orthogonal. That means that if we have two eigenfunctions ψ_m and ψ_n which belong to the different eigenvalues E_m and E_n, the following so-called orthogonality condition holds:

$$\int \psi_m \psi_n \, d\tau = 0 \tag{4-46}$$

To prove this we start from the Schrödinger equations for ψ_m and ψ_n

$$\nabla^2\psi_m + \frac{8\pi^2 m}{h^2}(E_m - U)\psi_m = 0 \tag{4-47}$$

$$\nabla^2\psi_n + \frac{8\pi^2 m}{h^2}(E_n - U)\psi_n = 0 \tag{4-48}$$

We multiply the first by ψ_n and the second by ψ_m and subtract the second from the first. By integrating over all space we get

$$\int (\psi_n \nabla^2 \psi_m - \psi_m \nabla^2 \psi_n)\, d\tau = \frac{8\pi^2 m}{h^2} (E_m - E_n) \int \psi_m \psi_n \, d\tau \qquad (4\text{-}49)$$

We transform the volume integral on the left by Green's theorem to a surface integral taken over an infinitely extended sphere

$$\int (\psi_n \nabla^2 \psi_m - \psi_m \nabla^2 \psi_n)\, d\tau = \oint (\psi_n \operatorname{grad} \psi_m - \psi_m \operatorname{grad} \psi_n)\, df \qquad (4\text{-}50)$$

which vanishes because the eigenfunctions, according to page 208, vanish at infinity. Therefore, it follows from (4-49)

$$(E_n - E_m) \int \psi_m \psi_n \, d\tau = 0 \qquad (4\text{-}51)$$

Thus the orthogonality condition is proved since it was assumed that $E_m \neq E_n$. We shall not consider here the degenerate case, $E_m = E_n$.

So far we have discussed only stationary states which are constant with respect to time and are described by the Schrödinger equation (4-42). In our derivation the time dependency of Ψ had been eliminated by expression (4-40). However, for many problems of atomic physics, e.g., collision processes, a time-dependent Schrödinger equation is needed. We obtain it by eliminating from (4-42) the eigenvalue of the energy E which is characteristic for a stationary state. By substituting in (4-40) $\nu = E/h$ and differentiating with respect to time, we obtain

$$\frac{\partial \Psi}{\partial t} = -2\pi i \frac{E}{h} \Psi \qquad (4\text{-}52)$$

or

$$E = -\frac{h}{2\pi i} \frac{1}{\Psi} \frac{\partial \Psi}{\partial t} \qquad (4\text{-}53)$$

By substituting this in the space-dependent Schrödinger equation (4-42) we obtain the complete time-dependent Schrödinger equation

$$\nabla^2 \Psi - \frac{8\pi^2 m}{h^2} U \Psi + \frac{4\pi i m}{h} \frac{\partial \Psi}{\partial t} = 0 \qquad (4\text{-}54)$$

The application of this theory to practical problems will be discussed in Sec. 4-7.

We now want to point out briefly the difference between the classical mechanical and the wave-mechanical treatment of atomic problems. Classical mechanics always proceeds directly or indirectly from the fundamental Newtonian equation

$$F = m \frac{d^2 r}{dt^2} \qquad (4\text{-}55)$$

and by integration arrives at the orbit of the considered point mass. This classical orbital motion has no place in wave mechanics, because the simultaneous exact knowledge of position and velocity of an atomic particle, which is a fundamental concomitant of the classical theory, is in contradiction to the uncertainty relation (page 198). Therefore, in atomic physics the Schrödinger equation (4-42) of wave mechanics replaces the basic Newtonian equation (4-55). In it the mass of the particle still occurs in connection with its potential energy [as in Eq. (4-55) in combination with the force]. However, in the result (solution) *the mass no longer appears, and the place of a mass point revolving in a closed orbit is taken by a stationary wave. Thus from wave mechanics (and even more radically from matrix mechanics to be discussed below) those quantities which are unsuited for describing atomic phenomena because they cannot be observed simultaneously have been eliminated.* The unsuitable classical description has been replaced by a description of the behavior of the physical system by a Ψ-function, whose physical meaning we shall discuss in the next section.

Before taking that up, we indicate in a few words the method used by Heisenberg, Born, and Jordan to solve the difficulties of the Bohr theory. While Schrödinger arrived at wave mechanics by developing the ideas of de Broglie on material waves, Heisenberg realized that the difficulties of the Bohr theory were especially due to unscrupulous application of such classical physical concepts to atomic problems, whose experimental determination, as we know, is fundamentally impossible. Heisenberg therefore radically rejected for his new quantum mechanics the use of any physical quantities which could not be measured by an experiment. Among the rejected quantities are, for instance, the use of the potential energy U (as in the simultaneous theory of Schrödinger), because it is derived by Coulomb's law from the assumption of point-like nuclei and electrons. What we know for certain about an atom are the frequencies of its spectrum and the intensities of its spectral lines. These quantities Heisenberg considered as given and sought to deduce all other characteristics of atomic systems and processes from them. The theory built on this basis does not use continuously variable quantities such as coordinates, but matrices consisting of discrete numbers, such as the frequencies of line spectra. It is called matrix mechanics. With the help of the principles of matrix mathematics developed earlier by mathematicians, an atomic theory was developed which is equivalent to wave mechanics in its completeness, but undoubtedly clearer and less open to principal objections. Certain basic problems such as transition probabilities and selection rules can be treated much easier and clearer by matrix mechanics than by wave mechanics. For the practical treatment of problems of atomic physics, on the other hand, it is usually simpler to use wave mechanics, because computations with differential equations are more familiar and simpler than those with matrices. For this reason it is of interest to note that the results of both theories not only agree quantitatively but that, in spite of their different starting points, wave and matrix mechanics are mathematically equivalent, as can be proved rigorously. In the following we shall use wave mechanics exclusively and thus need not go into matrix mathematics.

4-6. The Significance of Wave-mechanical Concepts. Eigenfunctions and Quantum Numbers. Spectral Intensity and Transition Probability

Now that we have laid the groundwork of the wave-mechanical formalism, we shall investigate the physical meaning of certain of the expressions under consideration. We have already learned *that the eigenvalues E of the Schrödinger equation are stationary energy states of an atomic system which is characterized by its potential energy U.* Although the significance of the eigenfunction ψ was not discussed, we did establish the fact that our interpretation must begin, not with the generally complex eigenfunction ψ itself, but with the norm of the eigenfunction, $\psi\psi^* = |\psi|^2$, obtained by multiplying ψ by its complex conjugate quantity ψ^*. $\psi\psi^*$ is always a real quantity.

Schrödinger interpreted the norm of the eigenfunction as the particle density at the point x, y, z and thus gave a very pictorial representation to the normalizing condition (4-45). By integrating the particle density over all space we must naturally get the whole particle. By applying the Schrödinger explanation to atomic electrons one is inclined, on the basis of this description, to speak of a "spread-out electron." The spatial representations of the eigenfunctions of electrons (or their norms) can then be regarded as *three-dimensional electron clouds* (see Fig. 116, page 227). This concept has the great advantage of being easily pictured and is frequently used, but it must be kept in mind that it is only a picture and is correct only as an average over some definite time interval.

The Schrödinger interpretation contains a number of fundamental inconsistencies. First, the mathematical treatment leads to the result that the eigenfunction, or its norm, for a system of more than one electron does not directly describe the behavior of the system in our three-dimensional physical space. A system containing N electrons actually is represented by its eigenfunction in a $3N$-dimensional space, the so-called configuration space, so that the explanation of $\psi\psi^*$ as a particle density in space is possible only by inference from a one-electron system. Second, an atomic system can be described simultaneously by several eigenfunctions, which, according to the Schrödinger interpretation, leads to the extraordinary conclusion that a particle is, or can be, in two different states at the same time. Third, we shall see (page 235) that in the Schrödinger sense a wave, representing a particle, can be split into several fractional waves by reflection and diffraction, which would mean that a particle could be broken into several parts. Fourth, according to this interpretation the wave-mechanical method itself is inconsistent, since the potential U in the Schrödinger equation is deduced from the assumption of point-like electrons and nuclei while, according to the interpretation of $\psi\psi^*$, their charges are distributed in space.

Born avoided these difficulties. Instead of Schrödinger's explanation, he proposed a statistical interpretation according to which the norm $\psi\psi^*$ *of an eigenfunction is the probability with which the particle in question is to be found by an experiment in a definite elementary volume.* According to this conception, the normalizing condition (4-45) simply means that the probability that a particle is somewhere in space must be equal to unity. The above objections are removed by this interpretation. The description of a system by *two* eigenfunctions means that in an experiment a certain percentage of the particles would be found in one state and the remainder in another. The separation of a Ψ-wave into several fractional waves by reflection and refraction at the interface of two media means that the probability of finding a particle on one side or the other of the interface is given by the norm of the eigenfunction of the fractional waves under consideration. The calculation of the potential energy U of the system, using the assumption of point charges of the nuclei and electrons, is completely in order since the particles have to be considered as points. This statistical interpretation of wave mechanics make the correlation with the uncertainty principle, page 198, especially clear. Instead of an exact statement about the actual location and the simultaneous velocity of a particle, which is not possible according to the uncertainty principle, we have in the statistical approach an expression for the probability of the behavior of the particle.

Now what is the relation of the description of the behavior of an atomic electron by the norm of its eigenfunction to that using the four quantum numbers n, l, m, and s? We have described these in Chap. 3 as representing what might be called the size, shape, orientation, and spin of the atomic electron. We have already mentioned that the electron spin is not yet included in pre-Dirac quantum mechanics, but has to be added to the wave-mechanical results afterwards. However, the other three quantum numbers must be contained in the eigenfunction ψ, because it is, in the Schrödinger sense, the solution of the problem of the atomic electron and contains *all* possible statements about it. We know that the norm of the eigenfunction describes the "vibration form" of the electron (compare the mechanical analogy, page 208). This form is characterized, just as in the mechanical case of the elastic vibration of a string, membrane, or elastic sphere, by the arrangement of stationary points, lines, or surfaces, the nodes of vibration. A circular membrane has *nodal curves* which are concentric circles and radii. A three-dimensional vibration such as the ψ-vibration consequently has three series of *nodal surfaces, namely nodal spheres (spherical shells), nodal planes going through the central point, and double conical nodal surfaces. These three series of nodal surfaces of the ψ-function, which characterize the modes of vibration of the standing ψ-wave in the atom, correspond to the three Bohr-Sommerfeld quantum numbers n, l, and m. In fact, the sum of all nodal surfaces is equal to $n - 1$, the number of nodal spherical surfaces equals $n - l - 1$, that of the nodal conical surfaces equals $l - | m |$, and that of the nodal planes equals m.* All this will become clear when we treat the theory of the H atom. Knowing this, we can easily translate electron symbols into the language of wave mechanics. For example, a $4p$ electron with $m = 0$ is described in wave mechanics by an eigenfunction which has two nodal spheres and one nodal cone but no nodal planes. The relationship of the old Bohr-Sommerfeld quantum symbols with the eigenfunctions of quantum mechanics is thus made clear. We will become acquainted with examples later on.

We now return once more to the Schrödinger interpretation of eigenfunctions and consider the wave-mechanical theory of radiation which stems from it.

For this we begin with the complete wave function Ψ which [see Eq. (4-40)], includes the time dependency, and keep in mind that, according to Schrödinger, the quantity

$$\rho = e\Psi\Psi^* = e\psi\psi^* \qquad (4\text{-}56)$$

represents the spatial charge density distribution (according to the statistical theory only its time average). This charge density distribution does not vary with time, because that part of Ψ which contains the time, $e^{-2\pi i\nu t}$, is eliminated by multiplying Ψ by its complex conjugate Ψ^*. According to electrodynamics, energy can be radiated only by an electric moment changing periodically with time. However, (4-56) proves that the charge distribution of the atom is constant with respect to time. As was pointed out on page 206, we thus have the exact wave-mechanical explanation of the absence of radiation by the stationary atomic states.

Just as an elastic string can vibrate simultaneously with two different frequencies whose amplitudes are superimposed, it should be possible also for a Schrödinger Ψ-function to be composed of several vibration forms, i.e., eigenfunctions. We will now obtain an expression for the charge density corresponding to Eq. (4-56), for the general case in which Ψ and Ψ^* do not belong to the same eigenfunction, but where the two eigenfunctions, Ψ_m and Ψ_n^*, are such that their superposition gives a true description of the vibration of the electron. Then we have

$$\Psi_m = \psi_m^0 e^{-2\pi i \nu_m t} \qquad \text{and} \qquad \Psi_n^* = \psi_n^{0*} e^{+2\pi i \nu_n t} \tag{4-57}$$

from which the charge-density distribution ρ_{mn} is

$$\rho_{mn} = e\psi_m^0 \psi_n^{*0} e^{2\pi i(\nu_n - \nu_m)t} \tag{4-58}$$

If the electron vibrates simultaneously in both eigenfunctions, Ψ_m and Ψ_n, we no longer have a charge density distribution which is constant in time, but a charge density that vibrates with a beat frequency $\nu_n - \nu_m$, which, according to classical electrodynamics, is associated with an emission of radiation of the same frequency. We replace the frequencies in (4-58) by the energy differences corresponding to them according to Bohr's frequency relation

$$\nu_n - \nu_m = \frac{E_n - E_m}{h} \tag{4-59}$$

We then see that *in wave mechanics the emission of a photon of frequency* $\nu_n - \nu_m = \nu_{nm}$ *is a logical consequence of an energy change* $E_n - E_m$ *of the atom, whereas in the Bohr theory it had to be introduced as a separate postulate (page 85).* In wave mechanics the beat frequency with which the charge density varies is equal to the frequency of the emitted radiation; no modification of classical electrodynamics is necessary. The Bohr transition from one energy state to another corresponds, in wave mechanics, to the simultaneous excitation of the two "vibration forms" (eigenfunctions) corresponding to these energy values. Thus wave mechanics encounters no formal difficulties in explaining the emission of radiation by atoms.

Moreover, there are eigenfunctions ψ_n and ψ_m the product of which, because of their special form, is zero. The same applies then to (4-58). In such cases the simultaneous existence of two such eigenfunctions of an atomic electron does not result in any radiation of energy. According to Bohr's theory the transition between the corresponding energy states is forbidden. *It can be shown exactly that in this way all selection rules which appear to be so arbitrarily stated in Chap. 3 are results of the special forms of the corresponding wave-mechanical eigenfunctions of the electrons.* Moreover, we can now understand how the selection rules can be made more or less invalid, for instance, by strong electric fields. The ψ-vibration of the electron and thus its eigenfunctions can be so altered (in a pictorial expression, "deformed") by the electric field that the product $\psi_n\psi_m^*$ in (4-58) which represents the intensity of radiation is no longer zero. When integrated over all space the result is at least a small positive time-dependent value of (4-58).

We define the transition probability A_{nm} as the probability (per unit of time) that an atom in state E_n spontaneously goes over into state E_m and emits a photon. Then the transition probability multiplied by the number of atoms in state E_n (page 166) determines the intensity of the spectral line emitted during the transition $E_n \rightarrow E_m$.

In order to compute spectral line intensities, we need (in agreement with classical physics) the components of the dipole vibration (4-58) along the x, y, and z axes. If we denote by $d\tau$ the volume element, we get from (4-58) for these so-called matrix elements the expressions

$$\left.\begin{aligned} X_{nm} &= e\int \psi_n x \psi_m^* \, d\tau \\ Y_{nm} &= e\int \psi_n y \psi_m^* \, d\tau \\ Z_{nm} &= e\int \psi_n z \psi_m^* \, d\tau \end{aligned}\right\} \tag{4-60}$$

From these matrix elements (time-independent components of the resultant dipole moment) we get the transition probability for a spontaneous transition $E_n \rightarrow E_m$,

$$A_{nm} = \frac{64\pi^4}{3h^4c^3} (E_n - E_m)^3 (X_{nm}^2 + Y_{nm}^2 + Z_{nm}^2) \tag{4-61}$$

This formula allows the computation of the radiation in any direction, i.e., including the polarization phenomena. For instance, if only one of the matrix elements (4-60) is different from zero, we have linear polarization, etc. We mentioned already that all the selection rules which can be derived from (4-60) and (4-61) follow with surprising simplicity from Heisenberg's matrix mechanics (page 211), from which the term "matrix element" has been transferred into wave mechanics. As a first example of his radiation theory, Schrödinger himself computed the intensities and polarization of the different Stark effect components of the Balmer lines of hydrogen (page 162) from (4-60) and (4-61) and found them to be in excellent agreement with the experimental results.

Thus the necessity no longer exists for using the correspondence principle, page 164, in wave mechanics.

This wave-mechanical theory of radiation, even though it seems so clear and has proved to be so successful, has the same defects which we mentioned in discussing the meaning of the Ψ-function on page 212. Moreover, the radiation theory based on the simultaneous excitation of the two eigenfunctions Ψ_n and Ψ_m cannot explain how the radiation actually stops after the excitation energy of the atoms is emitted, i.e., how all atoms eventually arrive at the stationary, nonradiating state E_m with the eigenfunction ψ_m. Dirac succeeded in eliminating this difficulty by treating in detail the interaction between the electric field of the electron vibrating in the Schrödinger manner, and the alternating electromagnetic field of the emitted or absorbed lightwave (or photon). He was able to show quantitatively how a change in the energy state of the atom, because of the corresponding change in the electric field effecting the binding between the nucleus and electrons, leads to the emission of electromagnetic radiation, i.e., photons, and conversely. However, this Dirac theory implied a quantization also of the electromagnetic field and this led to the beginning of quantum electrodynamics. This radiation theory of Dirac is able to describe quantitatively all radiation and emission processes in atomic systems. It does not have, on the other hand, the simple clarity of the first Schrödinger representation, which we discussed above.

4-7. Examples of the Wave-mechanical Treatment of Atomic Systems

After having become familiar with the fundamental facts of the formalism and meaning of wave mechanics, we now take up a few practical examples of the wave-mechanical treatment of atomic systems. Two of these examples, the rotator and the oscillator, are of especial interest in molecular physics, Chap. 6. The third example, which because of its complexity is treated last, is the H atom which serves as a test example in any atomic theory. Also in the case of these examples we shall concern ourselves less with the purely mathematical difficulties than with the fundamentally interesting physical questions, especially the reasons for the occurrence of the quantum conditions.

a. The Rotator with Rigid Axis

We begin our discussion with the simplest model for a diatomic molecule, the often discussed rotator with a fixed axis. It consists of two equal masses $M/2$, which are at a fixed distance from each other and which rotate about an axis perpendicular to the line joining the two masses (Fig. 113). This is also called the dumbbell model. This system has no potential energy; its entire energy is kinetic. Thus the Schrödinger equation (4-42) reduces to

$$\nabla^2\psi + \frac{8\pi^2M}{h^2}E\psi = 0 \tag{4-62}$$

Since we are dealing with rotation, it is appropriate to use polar coordinates. If we assume the rotator to be fixed in space, then the only variable is the angular coordinate of the rotation. Thus (4-62) can be written

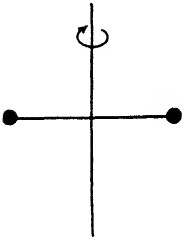


FIG. 113. Model of the rigid rotator, used as the simplest model for treating the rotation of diatomic molecules (see page 388).

$$\frac{1}{r^2}\frac{d^2\psi}{d\varphi^2} + \frac{8\pi^2M}{h^2}E\psi = 0 \tag{4-63}$$

By using the substitution

$$\frac{h}{8\pi^2cMr^2} = \frac{h}{8\pi^2cI} = B \tag{4-64}$$

where I is the moment of inertia of the molecule, we have the Schrödinger equation

$$\frac{d^2\psi}{d\varphi^2} + \frac{E}{hcB}\psi = 0 \tag{4-65}$$

The solution of this differential equation has the form

$$\psi = e^{im\varphi} \tag{4-66}$$

with

$$m^2 = \frac{E}{hcB} = \frac{8\pi^2Mr^2E}{h^2} \tag{4-67}$$

As the solutions ψ for the stationary states of the rotating system must be single-valued, ψ must have the same value after each complete revolution. That means we must have

$$e^{im\varphi} = e^{im(\varphi+2n\pi)} \qquad n = 1, 2, 3, \ldots \tag{4-68}$$

Thus m must be an integral number. From (4-67) it can be seen that the energy values of the rigid rotator are

$$E = hcBm^2 \qquad m = 1, 2, 3, \ldots \tag{4-69}$$

We thus have arrived at a very important result: because the solutions (4-66) *of the Schrödinger equation* (4-63) *must be single-valued* [*condition* (4-68)!], *the energy E of the rigid rotator cannot have any arbitrary value, but only the discrete eigenvalues given by* (4-69). *This is in contrast to the linear translational motion of electrons which is described by the Schrödinger equation* (4-44), *and to which no analogous condition for* *single-valuedness of the solutions applies. The quantization of the rigid rotator thus is a consequence of the obvious requirement* (4-68) *that the eigenfunctions must be single-valued.*

If we go from the rigid rotator to a rotator which is free in space, we must use in addition to φ a second angular coordinate, ϑ, which describes the position of the rotation axis with respect to the coordinate system. The simple Schrödinger equation (4-65) for the rigid rotator therefore takes on the more complicated form

$$\frac{1}{r^2}\left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\right)+\frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\varphi^2}\right]+\frac{8\pi^2M}{h^2}E\psi=0 \tag{4-70}$$

The somewhat wearisome calculations eventually lead to the energy eigenvalues of the free rotator

$$E = hcBm(m+1) = hcB(m+\tfrac{1}{2})^2 + \text{const} \tag{4-71}$$

In contrast to the rigid rotator, the energy states of the free rotator are characterized by *half-integral quantum numbers.* These half-integral rotation quantum numbers had already been derived from the analysis of the rotational structure of the band spectra to be discussed on page 387. Their explanation was an insurmountable difficulty for the old quantum theory, but it automatically follows from the wave-mechanical treatment of the rotator with its axis free to move in space. By plotting the eigenfunctions (4-66) of the rotator, it can be seen that they have nodal lines passing through the center of mass. The number of these nodal lines is equal to the quantum number m of the corresponding energy state of the rotator. Thus the rotational quantum numbers have found a wave-mechanical explanation.

b. The Linear Harmonic Oscillator

The next example we discuss is the linear harmonic oscillator, which we shall again meet as the simplest model of the vibrating diatomic molecules, page 365. A mass M which may be constrained to move along the x axis, may be bound to its rest position, taken as the origin of the coordinates, by an elastic force proportional to the displacement

$$F = -kx \tag{4-72}$$

Then it will vibrate through the origin with a frequency

$$\nu_0 = \frac{1}{2\pi}\sqrt{\frac{k}{M}} \tag{4-73}$$

The potential energy, which is taken to be zero at its minimum, at the equilibrium position, because of $F = -dU/dx$ and (4-73) is

$$U = \tfrac{1}{2}kx^2 = 2\pi^2M\nu_0^2x^2 \tag{4-74}$$

so that the Schrödinger equation for a linear harmonic oscillator becomes

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 M}{h^2}(E - 2\pi^2 M\nu_0^2 x^2)\psi = 0 \qquad (4\text{-}75)$$

Following Schrödinger by substituting

$$\xi = 2\pi x\sqrt{\frac{M\nu_0}{h}} \qquad (4\text{-}76)$$

and

$$\frac{2E}{h\nu_0} = C \qquad (4\text{-}77)$$

we get the differential equation

$$\frac{d^2\psi}{d\xi^2} + (C - \xi^2)\psi = 0 \qquad (4\text{-}78)$$

Assuming the solution to be

$$\psi = e^{-\xi^2/2}H(\xi) \qquad (4\text{-}79)$$

we obtain for $H(\xi)$ the differential equation

$$\frac{d^2H}{d\xi^2} - 2\xi\frac{dH}{d\xi} + (C - 1)H = 0 \qquad (4\text{-}80)$$

Since ψ and thus $H(\xi)$ must be finite through all space, the solution of this equation, $H(\xi)$, must be a polynomial and not an infinite power series. This is possible only if [see Eq. (4-77)]

$$C = \frac{2E}{h\nu_0} = 2v + 1 \qquad v = 0, 1, 2, 3, \ldots \qquad (4\text{-}81)$$

From this, we immediately obtain the possible energy values of the linear harmonic oscillator

$$E \doteq h\nu_0(v + \tfrac{1}{2}) \qquad v = 0, 1, 2, 3, \ldots \qquad (4\text{-}82)$$

In this case, the quantization (4-82) *follows from the boundary condition that the solutions* ψ, *and because of* (4-79) *the Hermitian polynomials* $H(\xi)$ *also, must be finite throughout space.* Again we have half-integral quantum numbers in agreement with the results of band spectroscopy (page 375). Equation (4-82) shows, furthermore, that *even in the lowest possible energy state* ($v = 0$) *the oscillator has a vibration energy* $E_0 = h\nu_0/2$, *the so-called zero-point energy. This could not be explained by the old quantum theory.*

Because the oscillator is so important in molecular physics, we shall discuss the probability of finding the oscillating mass at different distances from the rest position, according to both the classical and wave-mechanical picture (Fig. 114). Classically, the probability of finding the point mass is greatest at the point where the oscillation reverses its direction, because the mass remains there longest; this is shown by the values of the classical amplitudes of vibration in Fig. 114 for the first six oscillation states of the harmonic oscillator. On the other hand, the quantum-mechanical probability density, according to page 213, is given by the norm of the eigenfunction $\psi\psi^*$ which, computed from (4-79), is also plotted in Fig. 114. We see that the number of maxima of $\psi\psi^*$ is equal to the quantum number v plus 1. We also see that quantum-mechanically a mass with small probability can be at a distance from the origin where it *cannot* be according to the classical theory. We can see especially that in the lowest vibration state $v = 0$ the probability is greatest that a mass will be found at the center, where classically the probability is smallest. With increasing v, on the other hand, the central maxima of $\psi\psi^*$ become less significant while the intensive outermost maxima approach more and more the classical oscillation reversal points. Thus the higher the excited vibration state, the smaller are the differences between the classical and quantum mechanical behavior of the vibrating system. This is in agreement with our general law (correspondence principle page 165) that for large quantum numbers quantum physics goes over into classical physics.

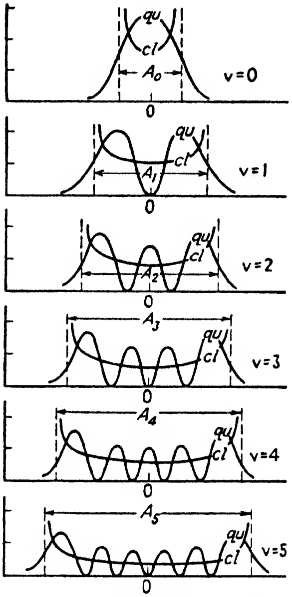


FIG. 114. Classical and wave-mechanical vibration of a harmonic oscillator. The probability of finding the oscillator at a certain distance from the rest position is plotted for the first six quantum-mechanically allowed vibration states; *cl* = according to the classical theory, *qu* = according to quantum mechanics. With a small probability the classically possible vibration amplitudes can be exceeded according to quantum mechanics as a consequence of the uncertainty principle. (*After Cl. Schäfer.*)

c. *The Hydrogen Atom*

We now discuss the famous wave-mechanical treatment of the H atom. It is not fundamentally more difficult, but it does require somewhat more detailed mathematics. This is due to the fact that we are dealing with a three-dimensional system in which all three polar coordinates are variable.

The potential energy of the proton-electron system is that of the Coulomb force of attraction

$$U = \frac{-e^2}{r} \tag{4-83}$$

so that the Schrödinger equation for the H atom is

$$\nabla^2\psi + \frac{8\pi^2 m_e}{h^2}\left(E + \frac{e^2}{r}\right)\psi = 0 \tag{4-84}$$

Written in the polar coordinates r, ϑ, and φ, (4-84) becomes

$$\frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r} + \frac{1}{r^2}\left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial\psi}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2\psi}{\partial\varphi^2}\right] + \frac{8\pi^2 m_e}{h^2}\left(E + \frac{e^2}{r}\right)\psi = 0 \tag{4-85}$$

For the solution we assume that ψ is the product of three functions, each of which is dependent on only one of the variables

$$\psi(r, \vartheta, \varphi) = R(r)\Theta(\vartheta)\Phi(\varphi) \tag{4-86}$$

Then (4-85) can be separated into three differential equations each of which is a function of only one coordinate.

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left[\frac{8\pi^2 m_e}{h^2}\left(E + \frac{e^2}{r}\right) - \frac{A}{r^2}\right]R = 0 \tag{4-87}$$

$$\frac{1}{\sin\vartheta}\frac{d}{d\vartheta}\left(\sin\vartheta\frac{d\Theta}{d\vartheta}\right) + \left(A - \frac{m^2}{\sin^2\vartheta}\right)\Theta = 0 \tag{4-88}$$

$$\frac{d^2\Phi}{d\varphi^2} + m^2\Phi = 0 \tag{4-89}$$

The constants A and m^2, the so-called separation constants, contain, as can be seen by comparison with (4-85), the constant terms in (4-87) to (4-88).

We recognize Eq. (4-89) as that of the rigid rotator [Eqs. (4-65) and (4-66)]; its solutions are, if we take into consideration the normalization Eq. (4-45),

$$\Phi = \frac{1}{\sqrt{2\pi}}e^{im\varphi} \qquad m = 0, \pm 1, \pm 2, \ldots \tag{4-90}$$

Equation (4-88) is not quite unknown to us either. It is what is left of Eq. (4-70) of the free rotator after (4-89) has been separated from it. Just as in the case of the fixed rotator, the solutions of (4-88), because they represent stationary states, must be single-valued, i.e., have the same value for ϑ as for $\vartheta + 2\pi$. From this condition it follows that, just as with (4-71), the quantization of Eq. (4-88) is of the form

$$A = l(l+1) \qquad l = 0, 1, 2, 3, \ldots \tag{4-91}$$

Introducing this into (4-88), we get

$$\frac{1}{\sin\vartheta}\frac{d}{d\vartheta}\left(\sin\vartheta\,\frac{d\Theta}{d\vartheta}\right) + \left[l(l+1) - \frac{m^2}{\sin^2\vartheta}\right]\Theta = 0 \tag{4-92}$$

The solutions of this equation, the associated spherical harmonics $P_l^{|m|}(\cos\vartheta)$ can be found in any mathematical table. In addition to (4-91) the secondary condition

$$l \geqq |m| \tag{4-93}$$

must be fulfilled.

The product of the solutions

$$\Phi(\varphi)\Theta(\vartheta) = e^{im\varphi}P_l^{|m|}(\cos\vartheta) \qquad \begin{matrix} m = 0, \pm 1, \pm 2, \ldots \\ l \geqq |m| \end{matrix} \tag{4-94}$$

gives the angular distribution of the eigenfunctions of the H atom.

We still have to solve the most important of the three separated equations (4-87) to (4-89), the r-dependent Eq. (4-87). It is the most important part of the original Schrödinger equation (4-85), because it alone contains the energy E which is characteristic for the atomic states. Making use of (4-91) we get from (4-87)

$$\frac{d^2R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left[\frac{8\pi^2 m_e}{h^2}\left(E + \frac{e^2}{r}\right) - l\,\frac{(l+1)}{r^2}\right]R = 0 \quad ✓\text{(4-95)}$$

This differential equation has single-valued solutions $R(r)$ which remain finite at infinity (boundary condition!), but only for certain discrete eigenvalues of the energy E. They are the energy states of the H atom (3-22) which we know already from the Bohr theory.

As the origin of our energy scale, we choose the state of ionization of the atom in which the proton and the separated electron are at rest relative to each other. With this stipulation the energy states of the bound electron have negative values (binding energy), while positive values of the energy E correspond to the kinetic energy of the ionized, free electron. In this energy scale, discussed on page 87, the electrons in elliptical orbits of the Bohr theory have negative values and those in the hyperbolic orbits (free electrons) have positive energy values (compare also page 106).

We treat next the solution of (4-95) for the stationary states $E < 0$. For large r's the terms with $1/r$ and $1/r^2$ in (4-95) can be neglected and we have

$$\frac{d^2R}{dr^2} + \frac{8\pi^2 m_e}{h^2}ER = 0 \tag{4-96}$$

Now if we substitute

$$\frac{h^2}{8\pi^2 m_e E} = -r_0^2 \qquad (4\text{-}97)$$

where r_0 turns out to be identical with the radius of the Bohr orbit of the ground state (page 97), and if we use for r a new variable which is twice the radius r measured in units of r_0,

$$\rho = \frac{2r}{r_0} \qquad (4\text{-}98)$$

we obtain the "asymptotic solution" of Eq. (4-96) for large r,

$$R = \text{const}\ e^{-\rho/2} \qquad (4\text{-}99)$$

For finite r or ρ we can replace the constant in (4-99) by a function of ρ which we shall consider later and write

$$R = e^{-\rho/2} w(\rho) \qquad (4\text{-}100)$$

By putting this expression in (4-95), we get for the equation containing $w(\rho)$

$$\frac{d^2w}{d\rho^2} + \left(\frac{2}{\rho} - 1\right)\frac{dw}{d\rho} + \left[\left(\frac{\pi e^2\sqrt{-2m_e}}{h\sqrt{E}} - 1\right)\frac{1}{\rho} - \frac{l(l+1)}{\rho^2}\right] w = 0 \qquad (4\text{-}101)$$

We try to represent $w(\rho)$ by a polynomial in order to be sure that R vanishes for large ρ (each polynomial for $\rho \to \infty$ approaches ∞ less rapidly than e^ρ). For this purpose, we set

$$w(\rho) = \rho^l u(\rho) = \rho^l \Sigma a_p \rho^p \qquad (4\text{-}102)$$

and use this expression and its derivative in Eq. (4-101) to obtain the recursion formula which, in order to remain finite, must break off at some particular term. This will occur for the pth term when

$$-(p + l + 1) = -n = \frac{\sqrt{-2m_e}\pi e^2}{h\sqrt{E}} \qquad (4\text{-}103)$$

From this the energy eigenvalues of the hydrogen atom are

$$E = -\frac{2\pi^2 m_e e^4}{h^2 n^2} \qquad \begin{array}{l} n = 1, 2, 3, \ldots \\ n \geqq l + 1 \end{array} \qquad (4\text{-}104)$$

By comparing this with (3-22) we see that this is exactly the equation for the energy values of the Bohr quantum orbits of the H atom, where n is the principal quantum number. *The quantum condition thus is a logical consequence of the boundary condition that the solution must be finite at infinity.* If we express the degree p of the polynomial $u(\rho)$ of (4-102) by n and l, we can write the solution

$$R(\rho) = e^{-\rho/2}\rho^l \underset{n-l-1}{u(\rho)} \qquad (4\text{-}105)$$

Thus we have the complete eigenfunctions of the H atom except for the complicated normalization factor from (4-94) and (4-105)

$$\psi = e^{-\rho/2}\rho^l \underset{n-l-1}{u(\rho)} P_l^{|m|}(\cos\vartheta)e^{im\varphi} \tag{4-106}$$

These eigenfunctions describe completely the behavior of the electron in the stationary energy state which is characterized by the three quantum numbers n, l, and m. Thus we have the following conditions for the quantum numbers

$$\left.\begin{aligned} n &\geqq l + 1 \\ l &\geqq |m| \geqq 0 \\ m &= 0, \pm 1, \pm 2, \pm 3, \ldots \end{aligned}\right\} \tag{4-107}$$

The fourth quantum number, describing the spin s, according to page 192, does not follow from the simple form of quantum mechanics used here; it must be added as a supplement.

We notice that in (4-104) the quantum numbers l and m do not occur, but only the principal quantum number n. Accordingly, n^2 different eigenfunctions are associated with any energy eigenvalue E_n, since l ranges from zero to $n - l$, and as to each l there exist $2l + 1$ orientations of l (i.e., different m values), and thus $2l + 1$ eigenfunctions,

$$\sum_{l=0}^{l=n-1}(2l + 1) = \frac{n}{2}(1 + 2n - 1) = n^2 \tag{4-108}$$

The energy eigenvalues of the H atom thus are $(n^2 - 1)$-fold degenerate (page 162). The degeneracy will be more or less removed by perturbations, e.g., as a result of the motion of the nucleus (page 98) and in the case of the Stark and Zeeman effect, page 155.

The r-dependent portions of the eigenfunctions for the first three states of the H atom are shown in Fig. 115; the small vertical line in each case indicates the radius of the corresponding Bohr orbit. Figure 116 is a pictorial representation of the norm $\psi\psi^*$ of the H atom eigenfunctions. The pictures show the value of the probability that an electron is at a point in space or, in other words, the time average of the electron density. These representations are very important for the understanding of chemical binding, especially in stereochemistry (see page 413).

With this we conclude the wave-mechanical treatment of the stationary states of the H atom ($E < 0$) and their eigenfunctions. We consider now briefly the states of positive energy ($E > 0$) which correspond to hyperbolic Bohr orbits and which are responsible for the continuous spectra (page 106) of the H atom. Since these states correspond to an electron wave proceeding from (or directed toward) the nucleus in the de Broglie sense, it is clear that there can be no quantum conditions. The wave or eigenfunctions of an electron beam proceeding from the nucleus with kinetic energy $E > 0$ are sine waves, as was discussed on page 208.

The wave-mechanical computations are carried out in the same way as for the case when $E < 0$, except that because of the positive sign of E we write, instead of (4-97)

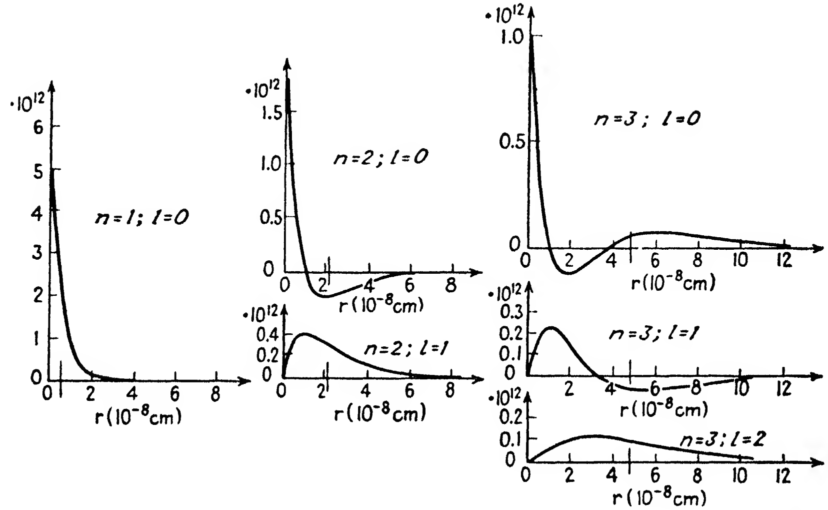


FIG. 115. The r-dependent parts of the electron eigenfunctions of the H atom for different values of the principal quantum number n and the orbital quantum number l. Note the different scales of the ordinates! The vertical marks on the abscissas indicate the radius of the corresponding Bohr orbits. (*After Herzberg.*)

$$\frac{h^2}{8\pi^2 m_e E} = +r_0^2 \tag{4-109}$$

Thus we have for the radial solution of (4-95) the expression

$$R = e^{\pm i\rho/2} w(\rho) \tag{4-110}$$

The double sign in the exponent means than an electron wave can move away from as well as toward the nucleus. The first case corresponds to the ionization of the atom, the second to the recombination of the electron and nucleus to form a neutral atom. By substituting (4-110) into the Schrödinger equation (4-95) we arrive at a differential equation similar to (4-101) for $w(\rho)$, but with complex coefficients, whose solution can again be represented by a power series. However, in contrast to (4-102) this series converges as $\rho \rightarrow \infty$ so that the solution is finite at infinity

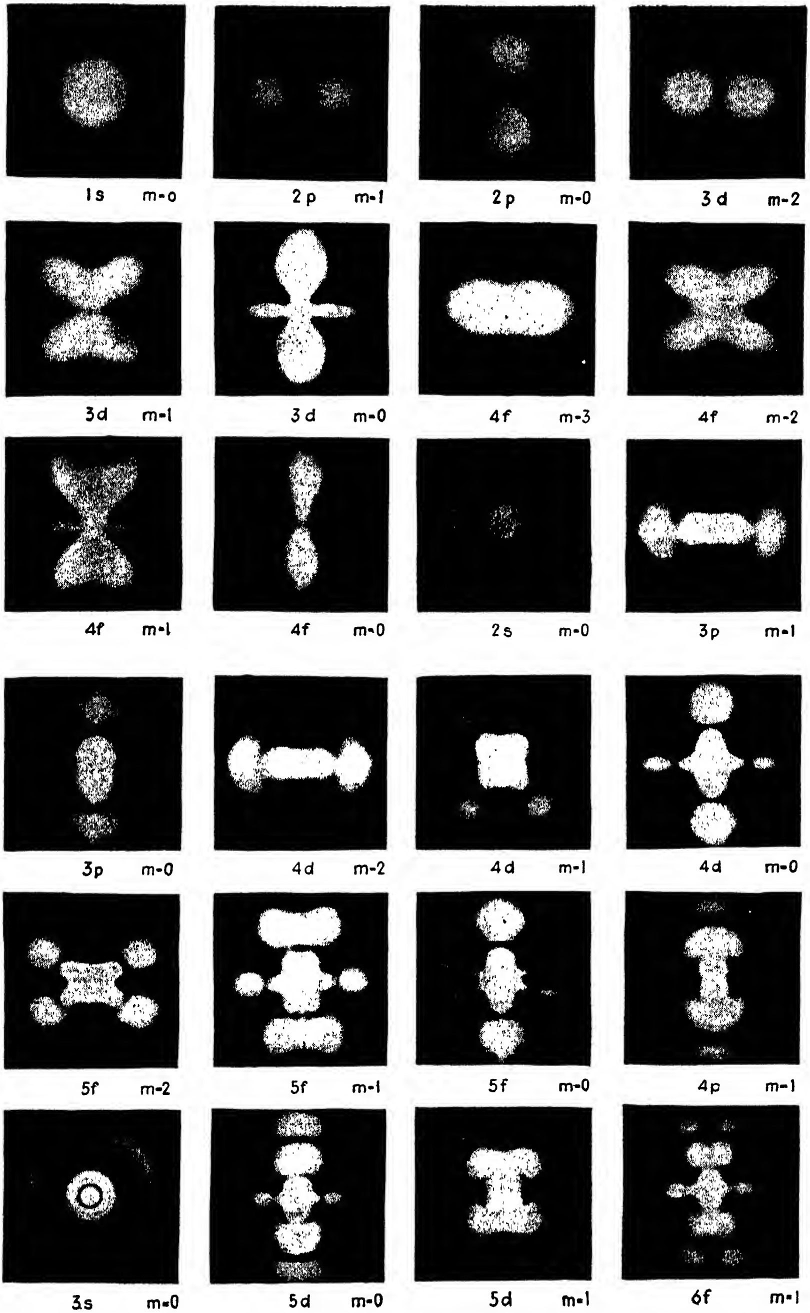


FIG. 116. Wave-mechanical "pictures" of the hydrogen atom in its different excited states (electron configurations). The brightness of light areas indicates the probability of finding the electron at that spot. (*Courtesy of H. E. White.*)

without the introduction of any special finiteness condition such as (4-103). However, since this finiteness condition resulted in the quantization of the E values for $E < 0$, we see from the mathematical formalism that there are no quantum conditions for $E > 0$. There are, on the contrary, for *all* positive E values solutions of the Schrödinger equation (4-95), which is in agreement with our nonmathematical discussions above. *It is one of the most satisfactory results of wave mechanics, that from the mathematics of the Schrödinger equation* (4-95) *we arrive at the discrete as well as the continuous energy values of the H atom corresponding to its line spectra and series limit continua, without making use of any arbitrary assumptions.*

Since the solutions for the free electron $E > 0$ are not stationary states but time-dependent waves traveling in either direction, we must consider the complete time-dependent wave function Ψ. The asymptotic solutions (for large r) of an electron with kinetic energy $E > 0$, therefore must be written

$$\Psi(r) = Ce^{i\left(\pm\frac{r}{r_0} - \frac{2\pi Et}{h}\right)} \tag{4-111}$$

corresponding to the expected progressive sine wave of frequency $\nu = E/h$ which is finite at infinity.

4-8. Interaction of Coupled Atomic Systems. Exchange Resonance and Exchange Energy

A very fundamental problem which was first solved by quantum mechanics is that of the energy resonance of coupled atomic systems of the same kind. It plays a very important role in the problem of the He atom (and all other many-electron atoms) because of the identity of electrons in the same discrete energy states (determination of multiplet spacing, page 142), and more so in the case of the interaction responsible for molecular binding of two identical atoms (page 410), in the case of the interaction of a large number of identical atoms or ions in a crystal (see page 438), and in the interaction of the external electrons of atoms of ferromagnetic materials where it is responsible for the phenomenon of ferromagnetism (page 449). The problem in all these cases is: *what happens if two atomic systems of equal total energy E are coupled with each other, i.e., if their interaction is taken into account?* Since these systems consist of charged elementary particles which are in motion, there always exists an interaction due to electric and magnetic fields.

In order to get an idea of the effects of coupling of two atomic systems, we consider the macromechanical case of the coupling of two pendulums.

However, we want to emphasize that we are not discussing two identical physical phenomena here, but only a mechanical analogy in order to clarify certain features of the quantum-mechanical phenomenon. Instead of the energy resonance of atomic systems we have, in the case of the pendulum, a mechanical frequency resonance. If we tie together two completely identical pendulums with a rubber band and set one to vibrating, the vibration energy is gradually transferred to the second pendulum until the first pendulum comes to rest. Then the energy is transferred back from the second to the first, and so on. Each of the two pendulums, which in the uncoupled state oscillate with a purely sinusoidal motion of constant amplitude (we disregard here any damping), oscillates in the coupled state as shown in Fig. 117. *As a consequence of the coupling of the two pendulums of equal frequency, beat oscillations appear. Such beat oscillations usually occur only when two slightly dissimilar frequencies are superimposed,* for example, by simultaneous excitation of two slightly different tuning forks. We thus can say: *by coupling two mechanical systems which in the uncoupled state have the same frequency, this original frequency splits into two different frequencies, one higher and one lower than the original one, and their difference is the greater, the stronger the coupling and, consequently, the interaction of the two systems.*

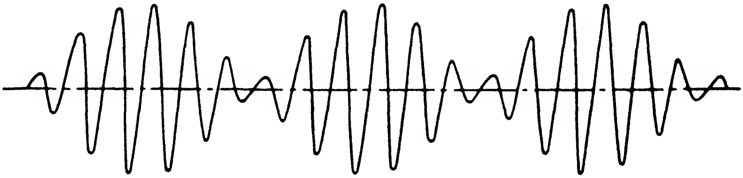


FIG. 117. Beat vibrations as a result of the coupling of two identical pendulums (schematic).

These results lead us to an understanding of the quantum-mechanical energy resonance. In order to examine the interaction of two identical H atoms in the ground state, we introduce in the Schrödinger equation the potential

$$U = -e^2\left(\frac{1}{r_{a1}} + \frac{1}{r_{b2}} + \frac{1}{r_{b1}} + \frac{1}{r_{a2}} - \frac{1}{r_{12}} - \frac{1}{r_{ab}}\right) \qquad (4\text{-}112)$$

r_{a1}, r_{a2}, r_{b1}, and r_{b2} are the distances of electrons 1 and 2 from the nuclei a and b, Fig. 118; r_{12} is the distance between the two electrons and r_{ab} the distance between the two nuclei. If we consider only the first two terms in the parentheses, which is the potential of electrons 1 and 2 with respect to "their own" nuclei, the Schrödinger equation separates into two equations for two uncoupled H atoms. The remaining terms of (4-112) are the interaction terms, and they represent the electrostatic potentials between the electrons and the nuclei which originally were not associated with them. In order to take into account these interaction terms in the Schrödinger equation, we must go through a complicated "perturbation computation." Its result is that the identical energy eigenvalue of the two uncoupled atoms splits, due to the interaction, into two eigenvalues, one of which is larger and the other smaller than the original energy state. The magnitude of the spacing is proportional to the magnitude of the interaction of the atoms.

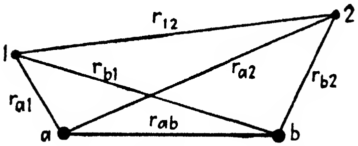


FIG. 118. Interaction of two hydrogen atoms in the H_2 molecule. a and b = hydrogen nuclei, 1 and 2 = electrons.

If we begin by considering the two H atoms without interaction, then the energy E of the total system

$$E = E_1 + E_2 = 2E_1 \quad (4\text{-}113)$$

evidently is equal to the sum of the energies of the two atoms in the ground state, while the eigenfunction of the system is equal to the product of the eigenfunctions of the individual electrons belonging to the nuclei a and b,

$$\psi' = \psi_a(1)\psi_b(2) \quad (4\text{-}114)$$

That ψ' is equal to the *product* of the individual eigenfunctions becomes clear when we consider that the eigenfunctions represent probability distributions. However, since the two electrons are equal in all respects,

$$\psi'' = \psi_a(2)\psi_b(1) \quad (4\text{-}115)$$

must *also* be an eigenfunction of the system. Since the two solutions ψ' and ψ'' belong to the same eigenvalue, we say that this eigenvalue and thus the system is degenerate. And since the two solutions result from an exchange of electrons 1 and 2, we speak of an *exchange degeneracy.* If we have a crystal composed of N equal atoms, then, as a result of the possibility of an exchange of all N atoms, neglecting the interaction, each energy eigenvalue is $(N - 1)$-fold degenerate. According to the theory of differential equations, not only are the eigenfunctions (4-114) and (4-115) solutions of the system, but also any linear combination of them

$$\psi = \alpha\psi' \pm \beta\psi'' \quad (4\text{-}116)$$

So far we have neglected the interaction of the two systems of particles. Now if we introduce the electrostatic interaction, that is, consider the last four terms of (4-112), then the properties of the entire system are changed, and we must add a perturbation term to the eigenfunctions (4-116) and the eigenvalue (4-113). It is very important to note that the electrostatic perturbation of the two atoms affects the electron orbits characterized by the eigenfunctions ψ' and ψ'' in *different* ways. This perturbation thus removes the exchange degeneracy. We therefore get, in analogy to the mechanical case of the coupled pendula, two different energy states E_s and E_a of the system.

We have to consider this difficult point in more detail because of its important consequences. We introduce into the Schrödinger equation not only the complete potential energy from (4-112) including the interaction terms, but also substitute in place of the eigenfunction ψ its actual perturbed value $\psi + \varphi$ (where φ is as yet undetermined). Similarly, we substitute the perturbed value of the energy $E_0 + \epsilon$ for the unperturbed eigenvalue E_0. Then we have an inhomogeneous Schrödinger equation which, instead of zero, has a perturbation term on the right side. Then, from the orthogonality and normalizing conditions (page 209) which apply to *all* eigenfunctions, we get the energy eigenvalues of the perturbed system

$$E_s = E_0 + e^2C - e^2A \qquad (4\text{-}117)$$
$$E_a = E_0 + e^2C + e^2A \qquad (4\text{-}118)$$

where e, as usual, is the elementary charge. The meaning of the quantities C and A will be explained below. First we note that in both equations a positive term is added to the unperturbed energy eigenvalue, that corresponds to the Coulomb interaction energy. Its constant is designated by C. There appears, furthermore, a term for the exchange energy, e^2A. This term, however, may have either a positive *or* a negative sign. As a result of this the single (but degenerate) eigenvalue E_0 now splits up into two values E_s and E_a, whose energy difference, the term spacing

$$\Delta E = 2e^2A \qquad (4\text{-}119)$$

depends only on the value of the exchange integral A.

We shall not go into the details of the derivation, but only state the results of the computation of C and A, for which we find

$$C = \int\left(\frac{1}{r_{ab}} - \frac{1}{r_{a2}} - \frac{1}{r_{b1}} + \frac{1}{r_{12}}\right)\psi_a^2(1)\psi_b^2(2)\,d\tau \qquad (4\text{-}120)$$

and

$$A = \int\left(\frac{1}{r_{ab}} - \frac{1}{r_{a1}} - \frac{1}{r_{b2}} + \frac{1}{r_{12}}\right)\psi_a(1)\psi_b(2)\psi_a(2)\psi_b(1)\,d\tau \qquad (4\text{-}121)$$

In (4-120) $\psi_a^2(1)$ and $\psi_b^2(2)$ are the probability densities which we designate by ρ_1 and ρ_2. Integrated over all space and multiplied by e they represent the total charges of the electrons 1 and 2, belonging to the nuclei a and b. We may therefore write for (4-120)

$$e^2C = \frac{e^2}{r_{ab}} - \int \frac{e^2\rho_2}{r_{a2}}\,d\tau_2 - \int \frac{e^2\rho_1}{r_{b1}}\,d\tau_1 + \int\int \frac{e^2\rho_1\rho_2}{r_{12}}\,d\tau_1\,d\tau_2 \qquad (4\text{-}122)$$

From (4-122) we see that the term C actually is that part of the interaction energy which is due to Coulomb forces, since the first term is the repulsion between the nuclei a and b, the second and third the attraction between nucleus a and electron 2 and between nucleus b and electron 1 respectively, and the last term is the repulsion between electrons 1 and 2. From (4-121) it can be seen that the exchange integral A is very similar to the Coulomb integral (4-120), except that instead of the actual electron densities $e\psi_a^2(1)$ and $e\psi_b^2(2)$ we have the "mixed terms" representing the electron exchange, $e\psi_a(1)\psi_a(2)$ and $e\psi_b(1)\psi_b(2)$.

The spacing of the term multiplets in the atomic spectra (see page 135), according to (4-119), depends on the magnitude of the exchange energy e^2A, the meaning of which can now be understood somewhat better. We shall meet numerous other important applications of this exchange effect in the course of the next chapters.

There are two different eigenfunctions which belong to the two eigenvalues of the energy (4-117) and (4-118). They describe the behavior of the system in the states under consideration and thus are adapted to the special perturbation (interaction). We say that they are "chosen" by the perturbation from the infinite number of possible solutions (eigenfunctions) (4-116). These eigenfunctions of the system (in our case of the H_2 molecule in the ground state) which we call ψ_s and ψ_a have been computed to be

$$\psi_s = \frac{1}{\sqrt{2 + 2J}}(\psi' + \psi'') \qquad (4\text{-}123)$$

$$\psi_a = \frac{1}{\sqrt{2 - 2J}}(\psi' - \psi'') \qquad (4\text{-}124)$$

J is that part of the exchange integral (4-121) which is independent of r,

$$J = \int \psi_a(1)\psi_b(2)\psi_a(2)\psi_b(1)\,d\tau \qquad (4\text{-}125)$$

The indices s and a with which we have designated the two eigenfunctions stand for "symmetric" and "antisymmetric," because ψ_s is symmetrical, i.e., does not change sign when the two electrons are exchanged, whereas the antisymmetric eigenfunction ψ_a in that case does change sign. The energy eigenvalues associated with these eigenfunctions bear the same indices.

As a consequence of the interaction, the originally degenerate energy eigenvalues of the coupled atomic system split up into a number of energy states equal to the number of particles which can be exchanged, whereby the magnitude of the energy spacing, the exchange energy, is dependent upon the value of the exchange integral (4-121). It is easy to see that this integral actually is a measure of the coupling of the two systems. The value of the integral is zero if in every volume element only *one* of the four eigenfunctions is zero. On the other hand, the value of the integral is the larger, the larger the values of all four eigenfunctions in every volume element, i.e., the more they overlap. If *one* eigenfunction vanishes at a point where the other ones are different from zero, it means that there is no coupling of the particle represented by this eigenfunction. Strong overlapping of the eigenfunctions, on the other hand, means that there is strong coupling. With this we conclude our discussion of exchange resonance. We return later to further applications.

In the previous discussion we did not take into consideration the spin of the electrons. We shall now do this because it is very important in many atomic problems. The electrons of two interacting atoms may have parallel or antiparallel spins, which means that both electron spins have the same or different signs, respectively. The Pauli principle, discussed on page 169, requires that in an atomic system no two electrons can be identical in all their properties (quantum numbers). This principle decides the question which of the two eigenfunctions (4-123) or (4-124), originating from the two interacting H atoms, applies to the normal H_2 molecule. If we have two electrons with the same n, l, and m in an atomic system, its eigenfunction *has* to change sign if we exchange those two electrons. Therefore, their spin quantum numbers must be $+\frac{1}{2}$ and $-\frac{1}{2}$, i.e., their spin directions have to be antiparallel. Expressed in other words, the Pauli principle requires that the eigenfunctions of electrons with the same m, l, and n be antisymmetric. Since in the stable ground state of the H_2 molecule, produced by the interaction of two H atoms, the two electrons must be in the lowest possible state, i.e., must be $1s$ electrons in the K shell, they must, according to the Pauli principle, have oppositely directed spins. Thus, without considering the spin, their eigenfunctions must be symmetric. Consequently, only the eigenfunction ψ_s can be the eigenfunction of the ground state of the H_2 molecule. Under consideration of the antiparallel spin directions of the electrons, ψ_s becomes antisymmetric with respect to the exchange of the electrons, in agreement with the Pauli principle. Parallel spin directions of the two electrons, on the other hand, must be characteristic for the state represented by the eigenfunction ψ_a, which is antisymmetric without consideration of the spin. The two electrons of the eigenfunction ψ_a having the same spin direction, consequently both *cannot* have the same quantum numbers, i.e., one of them must be excited. From this consideration it follows that the energy state E_a must lie above the state E_s. The computation of the perturbation, first carried out by Heitler and London, owing to the interaction of the two H atoms, Fig. 119, confirmed this result that only the eigenfunction ψ_s has a minimum of the potential energy for a definite distance of the nuclei, whereas the state E_a of the H_2 molecule, which belongs to the eigenfunction ψ_a, has no minimum of the potential energy and thus always lies above E_s.

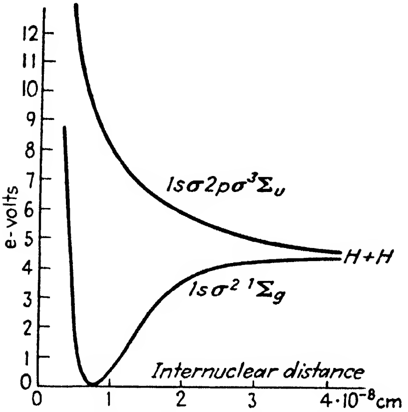


FIG. 119. Potential curves of the interaction of two H atoms after Heitler and London. The lower curve corresponds to the stable H_2 molecule (opposite spin directions of the two electrons). The upper curve (identical spin direction of the two electrons) corresponds to repulsion for every internuclear distance, i.e., mutual elastic deflection of the H atoms.

4-9. The Tunnel Effect (Penetration through a Potential Wall)

We are frequently concerned in atomic physics with the problem of whether or not a particle which is in a potential minimum (see Fig. 120) can surmount the surrounding potential wall and pass over to another potential minimum or to outer space. In classical physics the answer is simple: only if the kinetic energy of the particle is greater than the height U of the potential wall, can the particle get over the wall and "obtain its freedom." However, in the case of the radioactive α-decay it is evident that α-particles whose kinetic energy is much less than the height of the enclosing potential walls do leave the nucleus, page 275. Gamow and, independently, Condon and Gurney, were able to show that a quantum-mechanical treatment of this problem leads to entirely different results from those obtained from the classical theory. Quantum-mechanically, there is a definite probability that a particle can escape *through* the potential wall, though its kinetic energy is insufficient to go over the potential wall. This is called the quantum mechanical *tunnel effect.*

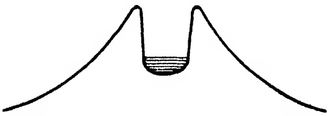


FIG. 120. Stationary states of the nucleons (protons and neutrons) in the potential well of a nucleus.

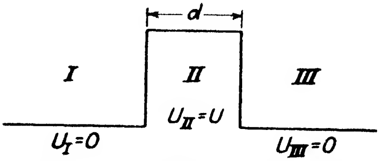


FIG. 121. Rectangular potential wall for the explanation of the tunnel effect.

We discuss the simplest case in which a particle, e.g., an electron, approaches a rectangular potential wall, of height U and width d, from the left, Fig. 121. If the potential on either side of the potential wall is zero, i.e., $U_{\text{I}} = U_{\text{III}} = 0$, and we assume that the particle is traveling in the x direction with a kinetic energy E, we can write for the Schrödinger equations in the three regions shown in Fig. 121

$$\frac{d^2\psi_{\text{I}}}{dx^2} + \frac{8\pi^2 m}{h^2} E\psi_{\text{I}} = 0 \tag{4-126}$$

$$\frac{d^2\psi_{\text{II}}}{dx^2} + \frac{8\pi^2 m}{h^2} (E - U)\psi_{\text{II}} = 0 \tag{4-127}$$

$$\frac{d^2\psi_{\text{III}}}{dx^2} + \frac{8\pi^2 m}{h^2} E\psi_{\text{III}} = 0 \tag{4-128}$$

The essential features of the problem are more easily understood by discussing an analogous case from the field of optics. We can see that according to Eq. (4-38), page 206, the phase velocity of the ψ-wave varies in regions of different potential U just as does the velocity of light in regions of different refractive index. From (4-38) we get $n = \sqrt{(E - U)/E}$. If we assume that the electron wave coming from the left in region I and incident on the surface I/II has a kinetic energy less than U, the wave will be totally reflected. Now we know from optics that when total reflection occurs at any surface, part of the light wave always penetrates this surface. However, its amplitude then decreases so rapidly (due to strong damping) that its depth of penetration is only of the order of a wavelength. This behavior of an optical wave in the case of total reflection is a consequence of the general continuity conditions at a surface. The same conditions hold for our electron wave. The electron wave incident on the surface I/II from the left will also be totally reflected, whereby an exponentially decreasing wave penetrates the potential wall II. If the width d is not much larger than the electron wavelength λ, computed from Eq. (4-26), the wave, with correspondingly smaller amplitude, can penetrate into region III, exactly as in the optical case of a thin lamina. Speaking in terms of a particle, we can say: in contrast to classical mechanics, there is a finite probability that an electron incident on the potential wall from the left can penetrate the potential wall and enter region III.

The complete wave-mechanical computations confirm these results. In addition to the Schrödinger equations (4-126) to (4-128) we need the continuity conditions for the surfaces I/II and II/III. They follow from the fact that in all three regions we are dealing with the same physically homogeneous Ψ-phenomenon. This means that at both surfaces the eigenfunctions ψ and their first derivatives $d\psi/dx$ must be continuous. With these conditions, there follows from Eqs. (4-126) to (4-128) that in region I we have, in addition to the incident wave, a reflected wave of constant amplitude, in region II a highly damped wave traveling to the right and reflected at the surface II/III, and in region III only a constant wave traveling to the right, but of very small amplitude compared to the wave in region I. *The norm of the particle wave in region* III *is the probability of the penetration of the potential wall by the particle.* This probability is shown by a fundamentally simple, but detailed, computation to be

$$\psi_{\mathrm{III}}\psi^*_{\mathrm{III}} = \frac{E(U - E)}{U^2} e^{-\frac{4\pi d}{h}\sqrt{2m(U-E)}} \tag{4-129}$$

The probability of a particle penetrating a potential wall is the larger, the smaller the width d and the smaller the height U − E of the potential wall, measured from the energy state E of the particle. This result, of which we shall find many applications, is one of the most important results of quantum mechanics. Though it is in direct contrast to the results expected from classical physics, it has been confirmed quantitatively by experiments in a most impressive way.

The tunnel effect can also be understood with the help of Heisenberg's uncertainty principle, page 200, Eq. (4-22). If the time Δt necessary to penetrate the potential wall is sufficiently small, the corresponding uncertainty of the energy ΔE is so large that there is a finite probability that E can exceed the height of the potential wall. From this presentation we can again see very clearly *that the character and significance of the uncertainty principle does not depend in any way on the accuracy of measurements.* We also see that our two expressions "penetrate" and, in the light of the uncertainty principle, "surmount" the potential wall, are merely two different ways of expressing the same physical process.

4-10. Fermi Quantum Statistics

In concluding our treatment of the more rigorous atomic theory, we discuss a topic of quantum physics which has proved to be of decisive significance especially for solid state physics, namely Fermi's quantum statistics. This theory is concerned with the energy distribution of electrons in an atomic system, for instance in a metal crystal.

Thus our question is, how does the energy distribution of electrons in a large atomic system depend on the absolute temperature, according to the classical Boltzmann statistics, and how does the quantum theory affect the energy distribution and its temperature dependence. We shall see that these changes are far-reaching and very fundamental.

As has already been emphasized in the introduction, we assume that the reader is sufficiently acquainted with classical Maxwell-Boltzmann statistics so that we have to consider only its basic features and consequences. Classical statistics is based on the assumption that the probability of a state, e.g., the probability that in our metal block just dN electrons have energies between E and $E + dE$, is determined by the number of possibilities by which this state can actually be realized. In our imagination we think of all particles (electrons) as numbered and determine how many different possibilities of distributing all N_0 electrons of the metal in the different energy cells of size dE there are which fulfill the requirement that there are just dN electrons in thermal equilibrium at absolute temperature T with energies between E and $E + dE$. The number of these possible distributions then gives the probability of the total distribution. By computing the number of electrons with energies between E and $E + dE$ as a function of T, we get the famous Maxwell energy distribution

$$dN = \frac{N_0 e^{-E/kT}}{\sum e^{-E/kT}}\, dE \tag{4-130}$$

According to this formula of classical statistics, at absolute zero of temperature all electrons have zero kinetic energy, whereas at higher temperatures the classical energy distribution (4-130) requires an exponential increase of the number of higher energy electrons.

What can be said about this energy distribution from the standpoint of quantum physics? In the first place, there is no continuous energy distribution possible in an atomic system, but only certain quantized energy states which are or can be occupied by electrons. For a metal crystal made up of N identical atoms, each energy state of the atoms which form the crystal splits up into N states as a result of the degeneracy caused by the interaction of the atoms. This fact was mentioned on page 230 and will be discussed in detail on page 438. According to the Pauli principle, each of these states can be occupied at the most by two electrons, which must differ in their spin directions, because otherwise there would be completely identical electrons of equal energy in the system. *Because of the existence of the quantized energy states and the validity of the Pauli principle, it is consequently* not *possible that all electrons have zero energy at the absolute zero of temperature. Moreover, quantum theory requires that they be distributed in such a manner over the lowest possible energy states of the crystal, that each state is occupied by only two electrons (of opposite spin). In contrast to the classical theory, consequently, at absolute zero those electrons which occupy the highest states have* considerable *kinetic energy, the so-called zero-point energy.* Its magnitude depends on the number of electrons per cubic centimeter of the crystal. The zero-point energy of the electrons in a metal is of the order of 10 ev. As a result of the validity of the quantum laws, metallic electrons thus have a kinetic energy at absolute zero which in the classical picture would require a temperature of about 100,000°K. Electrons in these states are called degenerate, and this phenomenon found by Fermi is called the *degeneracy of the electron gas.* It is of greatest importance for the electron theory of metals and thus for the theory of the electric conductivity of metals. It explains, for example, why electrons contribute comparatively little to the specific heat of a metal although they have to be regarded as practically free particles. Their zero-point energy is so great that small temperature changes produce only a minor change in the energy distribution of the electrons, though with increasing temperature gradually more and more electrons occupy energy levels which were vacant at $T = 0$.

It is also clear that the degeneracy of the electron gas in a metal vanishes as the average thermal energy of the electrons becomes of the same order of magnitude as the zero-point energy of the degenerate electrons. At very high temperatures, the distribution of the electrons over the energy states of the metal no longer is determined by the quantum laws or the Pauli principle, but by the temperature. This limiting temperature, above which classical statistics again becomes valid, is called the degeneration temperature T_0. It has been computed to be

$$T_0 = \frac{h^2}{8mk} n^{\frac{2}{3}} \tag{4-131}$$

where n is the electron density, m the electron mass, and k Boltzmann's constant. From Eq. (4-131) it follows that the degeneration temperature increases as the $\frac{2}{3}$ power of the electron density. Since in the common metals the electron density n is of the order of 10^{22} per cm^3, the degeneration temperature is of the order of 100,000°K, so that the electrons are always degenerate. On the other hand it follows from Eq. (4-131) that, for example, for arc discharges of very high electron density with $n \sim 10^{17}$ per cm^3, such as the high-current arc column investigated by the author, the degeneration temperature is of the order of only 100°K. The electrons of the arc plasma of about 10,000°K thus are not degenerate and do obey classical statistics. Also the phenomena of the electron gas degeneracy consequently can be explained in a satisfactory manner by the quantum theory.

From this presentation it is obvious that because of the validity of the quantum laws an entirely different statistical theory applies for the degenerate gas. We need it if we want to represent the energy distribution of degenerate electrons by a formula similar to (4-130). We now no longer are dealing with the possibilities of distributing electrons which can be distinguished from another. The problem is now only how the occupation of the quantized energy levels by basically indistinguishable electrons depends on the temperature. The derivation of this *Fermi statistics of degenerate electrons* leads to the energy distribution formula

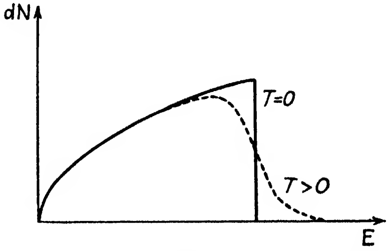


FIG. 122. Fermi distribution of the electrons in a metal. Ordinate = number of electrons, abscissa = electron energy. Solid curve for temperature $T = 0$, dotted curve for temperature $T > 0$ (schematic).

$$dN = C\frac{E^{\frac{1}{2}}}{e^{(E-kT_0)/kT}+1}\,dE \qquad (4\text{-}132)$$

where C is a constant, k again is the Boltzmann constant, and T_0 is the degeneration temperature determined from (4-131). To illustrate the energy distribution of electrons according to Fermi statistics, in Fig. 122 dN is plotted against the energy E for $T = 0$ and for a higher temperature $T > 0$. For $T = 0$ the number of electrons increases with the energy up to a certain limit, the energy of the highest occupied state, and is zero for all higher energies. At higher temperatures a fraction of the electrons, which at $T = 0$ occupy the highest filled levels, now occupy still higher levels which were not occupied at all at $T = 0$. We thus have the second energy distribution shown dotted in Fig. 122. For $T > 0$ we thus have a "tail" of fast electrons which is very important, for example, for the explanation of the thermionic emission of electrons from metals, page 446.

4-11. Achievements, Limitations, and Philosophical Significance of Quantum Mechanics

If we review quantum mechanics as discussed in this chapter and inquire about its achievements and limitations, we find that it describes practically the entire field of atomic, molecular, and solid state physics, as well as many essential processes of nuclear physics, and that it is in complete quantitative agreement with the experience. The progress which it has accomplished in many domains in which the old Bohr-Sommerfeld quantum theory failed is tremendous.

The difficulties concerning the arbitrary introduction of Bohr's postulates, which were necessary in order to explain the atomic spectra, are removed. The fact that atoms do not radiate in their stationary states was explained without violating the laws of electrodynamics. Quantum mechanics gave us, as logical consequences of the mathematical theory, not only the necessary quantum conditions for the energy levels, but it also explained the emission and absorption of radiation according to the Bohr frequency condition upon transition from one energy state to another, including the selection rules, the polarization of emitted spectral lines, and the line intensities. The anomalous Zeeman effect is explained just as quantitatively as are the finer features of the rotation and vibration of molecules including their interaction with the electronic shell. This could not be computed from the old theory (Chap. 6). A large number of problems which could not be understood at all by the old theory have been solved by means of the quantum-mechanical exchange theory (page 228). The problem of the helium atom and other atoms with several equivalent electrons belongs to this group, just as do homopolar binding and the energy bands of crystals with very many constituent atoms of the same kind (Chap. 7) or the phenomena of ferromagnetism which were so mysterious for many decades. Also a fundamental approach to the problem of nuclear forces was found on the basis of the exchange theory. The typically quantum-mechanical effect of the penetration of a potential wall by a particle (the tunnel effect, page 234) has made it possible to understand radioactive α-decay (page 278), the emission of electrons by a metal in a strong electric field (page 448), as well as the finer features of predissociation which will be discussed on page 380. The Smekal-Raman effect (see page 351), which was predicted on the basis of quantum physics and later verified experimentally, is of special importance in molecular physics. Finally, a fundamental achievement of quantum mechanics is that it explains the physical significance of the wave-particle dualism of light as well as of matter and, with the uncertainty principle, gives us a means of determining the kind of statements, and their accuracy, which may be made in the field of microphysics.

Quantum mechanics and its extensions (quantum electrodynamics) thus can be regarded as the correct theory for the quantitative description of phenomena and processes in the field of the atomic shells, including the interaction with radiation. Any statements and predictions made in these fields of physics on the basis of quantum-mechanical computations can be made with almost unlimited *confidence.* On the other hand, the limitations of the theory become evident. There are, in the first place, the difficulties which are associated with the unexplained "structure" of electrons and other elementary particles which can be regarded only in a first, coarse approximation as point charges. As will be shown in the next chapter, there are, furthermore, phenomena of nuclear physics which cannot be satisfactorily described by quantum mechanics as we have treated it here. Among these is the problem of the nuclear forces which hold the neutrons and protons together in the nucleus, a problem which is evidently related to the emission of electrons by the nucleus (β-decay, page 279), and to the π-mesons, to be discussed on page 332. It is possible, as we shall show in the next chapter, to draw an analogy between the nuclear forces and the quantum-mechanical exchange forces. However, it is evident that quantum mechanics in its present form does not give a satisfactory quantitative description of the nuclear forces. This failure of present-day quantum mechanics becomes especially clear in the case of meson physics, or such extreme nuclear processes as the generation of a large number of electrons, mesons, and possibly heavier elementary particles in a collision of *one* very energetic proton or neutron with a nucleus (see page 324). These multiple materialization processes cannot be explained by our present quantum mechanics. This requires, according to Heisenberg, an extension of the theory, which leads to the assumption of a smallest length l_0 of the order of 10^{-13} cm as a new universal constant, and a corresponding "smallest time," the time required to traverse the smallest length with the velocity of light, l_0/c. It seems that present-day quantum mechanics is suitable to describe only those atomic processes which involve no essential change in the state of the system within the smallest time. It is apparent that this condition is not fulfilled for such extreme processes as a multiple materialization of kinetic energy in an explosion shower. It is not yet clear what the improved quantum theory, toward which Heisenberg, Dirac, Schwinger, Dyson, Bopp, and others are striving, will look like.

Before concluding this chapter we must finally analyze, at least briefly, the fundamental difficulties and problems which have been raised by quantum mechanics and which are quite generally closely related to the philosophy of science. These conceptual difficulties of quantum physics undoubtedly are so large that even today many scientists who are mainly engaged in experiments (and so many philosophers!) regard quantum mechanics and its description of the world with undisguised skepticism — a much too complacent attitude since the great results of the theory which we have described, and their logical consequences, should have inspired them to a very serious analysis. Three objections are raised most frequently against the quantum-mechanical theory: its incomprehensibility, its indeterminacy, which is evidenced by the fact that it can be described only statistically, and, related to this, its asserted acausality.

Now there is, in fact, no doubt that our new physical "picture" of the world is much less pictorial than that of classical physics of the year 1900. The theory of relativity already had stripped our concepts of space and time of their seemingly so obvious absoluteness. It had changed the classical meaning of matter and energy greatly by teaching us that mass and energy can be converted into each other. In quantum mechanics, the fundamental concept of classical physics, force, has not only lost its predominant position but has actually been eliminated from the theory. The decisive role which Planck's elementary quantum of action h plays in all physics seems, moreover, to point very clearly to a dominant significance of the very unpictorial concept of action (energy times time) for all physics. Evidence for this central role of the concept *action* can also be derived from the relativity theory. Whereas the values of force, mass, and energy depend on the state of motion of the reference system, the action is the only physical quantity which is independent of the reference system, time and place. It is "relativistically invariant." Both quantum theory and relativity theory thus agree about the central significance of action in physics, so that *it seems appropriate to regard the quanta of action h as the last physical realities upon which our whole manifest world depends.*

We have seen on page 198 that the uncertainty principle and the wave-particle dualism are also based on this dominant role which the basically unpictorial concept of the quantized action plays in modern physics. And it is the uncertainty principle and the wave-particle dualism which cause the fundamental conceptual difficulties in our new "picture" of the world.

The fact that all atomic particles, depending on the experiments we perform with them, display such contradictory properties as those of a particle, well localized in space and time, or those of an extended wave field, has important consequences. One way out of these apparent contradictions is favored today by many physicists: the assumption that only the particles and not the waves are "real." In terms of the Born explanation of wave mechanics (page 213), waves are only probability waves which "guide" the particles. The wave pattern which we observe as a result of an electron diffraction experiment (Fig. 108*a*) shows us only the statistical distribution of the diffracted electron particles which were guided by the probability waves so as to form this pattern.

Does this explanation really remove the difficulties of the wave-particle dualism? Let us consider the diffraction of an electron beam by one or two slits. The diffraction pattern is different in both cases. Now we have to assume either that an electron which just passes one slit "knows" whether the second slit is closed or open (this seems an impossible assumption) or that the guiding waves (Ψ-waves) are different in both cases because their interference in both cases results in a different interference pattern. Then, however, the conclusion seems evident that the guiding waves are just as "real" as the guided particles, and we face again the dualism which we tried to evade, only in a somewhat different form. The Ψ-waves (probability waves) then describe the behavior of the corpuscular particles in their interaction with the experimental arrangement under consideration (e.g., with the one or two diffracting slits), but *only in a statistical way.* This description of atomic particles with their dependence on the potential field $U(x,y,z)$ in which they move, and on the boundary conditions given by the experiment, is typical for wave mechanics. It means that *we cannot make any statements about "atoms as such," but only about particles interacting with their surroundings.* However, because of the statistical relation between the particles and their guiding waves (in the Born explanation) we never can evade the consequences of the uncertainty principle, discussed on page 198. No matter how we look at it, the determinacy of all properties and events, so self-evident in the world of classical physics, does not apply to microphysics.

The consequences of this failure of complete determinacy in atomic physics have led to what is called the acausality of microphysics. Its actual meaning has been misunderstood frequently and has led, and leads, to untenable objections against quantum mechanics itself. The most serious of these objections is that the indeterminacy or acausality (both terms often used without clear analysis as to their actual meaning) would make any research on atomic problems impossible, because every measurement obviously is based on a causal connection between the event to be studied, and its effect on the measuring instrument. This objection fails to take into account the actual meaning of the quantum-mechanical indeterminacy. *If by interaction with some experimental arrangement we* *"force" the atom to reveal some of its properties, the exact relation between cause and effect applies to these properties and their dependence on space and time, just as in classical physics. However, according to quantum mechanics, in this case no statements can be made about those properties of the atomic system which are complementary to the measured ones in the sense of the uncertainty principle.* Consequently, with respect to a connected sequence of measurements, the law of causality remains valid also in microphysics, and there is no reason whatsoever to doubt the correctness of our experimental study of atomic phenomena.

The influence of the uncertainty principle thus causes a restriction of the applicability of the causality principle, especially a restriction of its thoughtless application. In classical physics causality is defined by the assertion: "If the state of a closed system is known completely and precisely at one moment, then it is possible, by principle, to compute the state of the system for any earlier or later time." It is obvious from the uncertainty principle that this classical causality principle is unapplicable, not because of any basic invalidity, but because its condition cannot be fulfilled. It is not possible to determine the state of any system at any definite time precisely and completely. C. F. von Weizsäcker has pointed out that also in classical physics actually always only certain properties are known, and conclusions concerning these properties are drawn, while the unknown properties are without influence. This limitation of the causality principle, which actually applies also to classical physics, has been recognized by quantum mechanics as a *fundamental* restriction. Von Weizsäcker, therefore, formulates the principle in the following way: "*If certain properties of a system are precisely known at a certain time, then all those properties of earlier or later states of the system can be computed which, according to classical physics, are causally connected with them.*"

This restriction of the causality principle has no bearing, consequently, on the possibility of measuring atomic phenomena or events. There is, however, a definite lack of determinacy, as far as single atomic events are concerned, again as a consequence of the uncertainty principle. We shall learn in the next chapter (page 273) that the exact moment at which a certain radium nucleus will emit an α-particle is absolutely undetermined. All we can find out is that after 1,590 years half of the initial number of Ra nuclei will have emitted an α-particle. The single event remains undetermined; the only thing we can determine, experimentally or theoretically, is a statistical law which refers to the behavior of a large number of particles. The same holds true for all atomic physics. *The single atomic event thus is basically acausal and is determined only in a statistical sense by the guiding probability waves.*

Only for the collective behavior are definite statements possible. We can speak directly of a statistical determinacy.

No doubt, all this is quite a change in our concepts of physics compared with the classical, rigid causality according to which the whole world's course was supposed to follow automatically and mechanically from a given initial state, like a gigantic, though very complicated, clockwork. We shall not discuss here the problem, which is of great interest to biologists and philosophers, whether and to what extent this indeterminacy of the single atomic event leaves room for a certain "freedom" in the course of the world, though this point plays quite a role in the consequences of Jordan's quantum biology. We want to point out, however, that *the indeterminacy in quantum mechanics has to be regarded as final in so far as no future extension of the atomic theory will be able to reestablish the complete determinacy.*

In concluding our discussion of the philosophical consequences of quantum mechanics we should like to emphasize the fact that we evidently run into conceptual difficulties as soon as, and whenever, the progress of physics leads us beyond the range of the world which is "normally" accessible to man. The first major case where theoretical results seemed irreconcilable with our normal concepts of physics, or even with apparently self-evident philosophical concepts, was encountered by the theory of relativity. Velocities of the order of the speed of light were just as inaccessible to man's experience before the age of physics as distances of millions of light years, where the apparently so self-evident Euclidian geometry seems to fail. The same applies, however, to the atomic dimensions of microphysics which are, in the other direction, beyond our direct reach. The phenomena in this region are, for all we know, described correctly by quantum mechanics, but this description leads to the conceptual and philosophical difficulties outlined above. Is it really incidental that we run into such difficulties whenever we go beyond the range of the world "given" to man? Could it not be that these conceptual and philosophical difficulties stem from the fact that all our elements ("tools") of thinking have been developed and derived from man's interaction (i.e., experience) with "his" region of the world? Our very tools of thinking, then, are not suited for the immensely expanded world of our modern science, and our attempt to "understand" this wider world in a pictorial way necessarily leads to conceptual difficulties. However, *quantum mechanics has taught us that it is possible to adapt our ability of conception to the new situation and thus has opened new realms to our thinking. This seems to be a definite and positive contribution of quantum mechanics to philosophy.* It seems possible that from this angle new light might fall also on Kant's "a priori" concepts, still considered an essential prerequisite of any science. No matter how all this may develop, further analytical and synthetical work on all the relations discussed in this chapter and on their consequences certainly should lead us to a more satisfactory scientific-philosophical picture of our world than was the old mechanistic picture of classical physics.

Literature

Born, M., and P. Jordan: Elementare Quantenmechanik, Springer-Verlag, Berlin, 1930.

Brillouin, L.: Quantenstatistik, Springer-Verlag, Berlin, 1936.

Condon, E. U., and P. M. Morse: Quantum Mechanics, McGraw-Hill Book Company, Inc., New York, 1929.

Dänzer, H.: Grundlagen der Quantenmechanik, T. Steinkopf, Leipzig, 1935.

Dirac, P. A. M.: The Fundamental Principles of Quantum Mechanics, 3d ed., Oxford University Press, New York, 1947.

Dushman, S.: The Elements of Quantum Mechanics, John Wiley & Sons, Inc., New York, 1938.

Flügge, S., and A. Krebs: Experimentelle Grundlagen der Wellenmechanik, T. Steinkopf, Leipzig, 1935.

Frenkel, J.: Wave Mechanics, Oxford University Press, New York, 1934-1936.

Glasstone, S.: Theoretical Chemistry, D. Van Nostrand Company, Inc., New York, 1944.

Heisenberg, W.: Die physikalischen Prinzipien der Quantenmechanik, 2d ed., S. Hirzel, Leipzig, 1943.

Heitler, W.: Quantum Theory of Radiation, Oxford University Press, New York, 1945.

Hellmann, H.: Einführung in die Quantenchemie, Deuticke, Leipzig, 1937.

Jordan, P.: Anschauliche Quantentheorie, Springer-Verlag, Berlin, 1936.

Kemble, E. C.: The Fundamental Principles of Quantum Mechanics, McGraw-Hill Book Company, Inc., New York, 1937.

Kramers, H. A.: Die Grundlagen der Quantentheorie, vol. 1/I, II, Hand- und Jahrbuch der Chemischen Physik, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1937.

March, A.: Natur und Naturerkenntnis, Springer-Verlag, Vienna, 1948.

Mott, N. F.: Wave Mechanics, Cambridge University Press, London, 1930.

Mott, N. F., and I. N. Sneddon, Wave Mechanics and its Applications, Oxford University Press, New York, 1948.

von Neumann, J.: Mathematische Grundlagen der Quantenmechanik, Springer-Verlag, Berlin, 1932.

Pauling, L., and E. B. Wilson, Jr.: Introduction to Quantum Mechanics, McGraw-Hill Book Company, Inc., New York, 1935.

Reichenbach, H.: Philosophic Foundations of Quantum Mechanics, University of California Press, Berkeley, 1944.

Rice, F. O., and E. Teller: Structure of Matter, John Wiley & Sons, Inc., New York, 1949.

Rojansky, V. B.: Introductory Quantum Mechanics, Prentice-Hall, Inc., New York, 1938.

Schäfer, Cl.: Quantentheorie, vol. III/2, Lehrbuch der Theoretischen Physik, W. de Gruyter, Leipzig, 1937.

Schiff, L. I.: Quantum Mechanics, McGraw-Hill Book Company, Inc., New York, 1949.

Schrödinger, E.: Collected Papers on Wave Mechanics, Blackie & Son, Ltd., Glasgow, 1928.

Sommerfeld, A.: Atombau und Spektrallinien, vol. II, 4th ed., Vieweg, Brunswick, 1949.

van der Waerden, B. L.: Die gruppentheoretische Methode in der Quantenmechanik, Springer-Verlag, Berlin, 1932.

von Weizsäcker, C. F.: Zum Weltbild der Physik, 3d ed., S. Hirzel, Leipzig, 1945.

Wentzel, G.: Einführung in die Quantentheorie der Wellenfelder, Deuticke, Vienna, 1943.

Weyl, H.: Gruppentheorie und Quantenmechanik, S. Hirzel, Leipzig, 1928.

Wigner, E.: Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren, Vieweg, Brunswick, 1931.

CHAPTER 5

NUCLEAR PHYSICS

5-1. The Relation of Nuclear Physics to General Atomic Physics

From the standpoint of a rigid systematical approach, a discussion of atomic physics should begin with a discussion of the nucleus, since the essential properties of the nucleus, e.g., its charge, determine, partly at least, the structure of the electron shells of the atom. Then atomic physics in the more restricted sense, molecular physics and physics of the solid state, would logically follow the treatment of nuclear physics. In presenting the pictorial Bohr atomic physics first, we have followed the historical development and, at the same time, have the advantage of proceeding from the simpler to the more difficult. Historically, up to about 1927, the first field to be studied and explained was the physics of the atomic shells. Then further development led, on the one hand, to a theory of molecules which are composed of a number of atoms as well as to a theory of larger atomic complexes (liquids and crystals) and, on the other hand, to nuclear physics. For this development it was necessary to have the knowledge of the Bohr atomic theory and its quantum-mechanical refinements which we discussed in the last chapter. The same applies to our presentation in this book. Energy states and the transitions between them, accompanied by radiation, occur in the nucleus (though with correspondingly larger amounts of energy) just as in the electron shells of the atom. It is impossible to understand essential processes such as the decay of nuclei or the exchange forces effecting the binding between the nuclear constitutents without a knowledge of quantum mechanics. It thus seems reasonable, both from the standpoint of an intelligible introduction as well as from that of the actual development of our science, to discuss nuclear physics at this point.

In our presentation, we shall discuss only briefly the experimental and technical questions and the abundance of the known nuclear reactions, collectively known as *nuclear chemistry*. For further details of these subjects the reader is referred to monographs listed among the references at the end of this chapter. However, we want to stress those phenomena and theories of nuclear physics which are of fundamental interest in relation to the elementary particles, their nature, and their interaction with radiation, for nuclear physics has given us an entirely new and deeper insight into the meaning of matter and energy and their interrelationship.

Although the essential features of the structure of the nucleus, as in the case of the atom, can be represented pictorially in a manner which is of help for the further development of nuclear physics, definite and quantitative statements about the behavior of the nucleus and about nuclear processes can be expected only from an exact theory. The unsolved problems of nuclear forces, mentioned in the last chapter, have already shown us the limits of present-day quantum mechanics, which fails in certain domains of nuclear physics, though it describes accurately all processes occurring in and with the electron shells. Therefore, the nuclear reactions in high energy collisions, such as are found and studied in cosmic radiation, are of special interest to us. New elementary particles and new processes, such as the conversion of matter to radiation, have been discovered during these studies. The inclusion of the basic processes of cosmic radiation in our discussion appears therefore to be unavoidable, the more so since we expect many important results in this field in the future. We shall refer to this again.

5-2. Experimental Nuclear Physics

In order to understand the results of nuclear physics it is necessary to make a brief survey of the experimental methods used to study the nucleus. We disregard here the experiments already discussed in Chap. 2 which furnished us with information about the general properties of nuclei, and deal with the methods which allow us to detect and study nuclear processes. In the following section we shall discuss the methods of accelerating nuclear particles which are necessary in order to excite the nucleus and to induce nuclear transformations.

Methods used to detect and measure natural or induced nuclear processes always depend on the detection and the measurement of the energy of nuclear particles (α-particles, protons, neutrons, and larger fragments) or nuclear radiation (γ-radiation, electron and positron radiation). Because of the high velocities or energies which prevail in nuclear physics, all charged particles ionize the matter which they penetrate. *This ionization is used to detect and also to measure the energy of the charged particles.* Being uncharged particles, neutrons cannot produce ionization. However, they can knock protons out of a hydrogen-containing substance, and the ionization by these secondary protons can serve as a means of detecting neutrons and measuring their energy. The detection of neutrons will be taken up in detail on page 291. Also the γ-radiation of the nuclei has no direct ionizing effect. However, when a γ-quantum is absorbed by the electron shell of an atom, a photoelectron is emitted, which does produce ionization and thus serves as a measure of the energy of γ-radiation.

The measurement of energy by ionization methods depends on the experimental fact that *each ion pair formed in air requires* 32 *ev energy.* Thus, from the number of ion pairs formed, the particle energy can be determined, provided the total ionization up to the complete stopping of the ionizing particle occurs within the apparatus used for the measurement.

In addition to these ionization methods, the details of which we shall consider later, there is a second method for measuring energies. It utilizes the deflection of the particles in a magnetic field, in combination with the fact that the traces of the deflected particles can be made visible by their own ionization. According to the formula which we used in discussing the mass spectrograph, page 46, a particle of charge e, mass m, and velocity v will be deflected by a magnetic field H perpendicular to the initial direction of the particle through a circular path whose radius of curvature is given by

$$R = \frac{mcv}{eH} \tag{5-1}$$

If the mass and charge of the particle are known, its velocity v and thus its energy $\frac{1}{2}mv^2$ [or, relativistically, $mc^2 = m_0c^2/\sqrt{1 - (v^2/c^2)}$] can be derived from its curved path in the magnetic field. Usually the *charge* on the particle is known. The α-particle carries two elementary charges and all other charged particles that are known up to now (proton, electron, meson) carry one elementary charge. The *sign of the charge* can be determined from the direction in which the particle is deflected by the magnetic field. The *mass of the particle*, on the other hand, can be determined, at least roughly, from the number of ion pairs produced per centimeter of its path. Since an α-particle forms about 50,000 ion pairs per centimeter of path, a proton about 10,000 and an electron about 100 (these are very rough figures, since the values depend on the velocity) these particles can be distinguished by the ionization which they cause.

The most important device for exact nuclear studies is the *cloud chamber* (Fig. 123), first built by C. T. R. Wilson. By a rapid expansion the atmosphere in the chamber, which must be made dustfree to remove any nuclei for condensation, becomes saturated with water vapor. A charged particle flying through this space will produce ions which act as condensation nuclei for water droplets. The path of the particle then can be seen as a streak of mist and, if strongly illuminated by an arc lamp, it can be photographed (Fig. 138). By using the rough estimate given above, the approximate mass of the ionizing particles can be determined from the droplet density, and their energy can be determined from the range, or total ionization, providing the particle is totally absorbed in the chamber. This does not always happen, for instance, with the very fast particles of cosmic radiation. In this case the whole cloud chamber is placed in a magnetic field and the velocity and energy is determined from the curved path of the particle, especially if it is photographed stereoscopically so that the path can be measured in three dimensions. Figure 148 is an example of a stereoscopic photograph of a nuclear process in a cloud chamber. The disadvantage of the cloud chamber, which is satisfactory in all other respects, is that it has only a short "active" time after each expansion ($\frac{1}{10}$ to $\frac{1}{100}$ sec) and thus is not suitable for continuously recording nuclear processes. The useful time can be increased to several seconds by slowly expanding the chamber. Experiments have also been conducted with *continuously working cloud chambers* in which a region of supersaturated vapor was maintained by chemical or thermodynamical means. However, these chambers have not found much use so far, first, because the disturbance of new paths by the falling droplets of the older paths complicates the observation of the new paths, and second, because the convection currents set up by the local heating due to condensation prevent the formation of sharp traces. For continuous recording of nuclear processes the ionization chamber and the Geiger-Müller counter are used, not to determine the exact path of the particle, but to answer the question when and with what energies particles are emitted by a nuclear sample or pass through a certain volume. In the ionization chamber the particle simply passes through the electric field of a condenser and ionizes the gas between its plates. The ions thus produced, whose number is proportional to the energy of the ionizing particle, are attracted to the

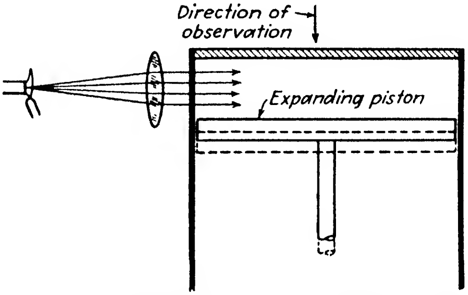


FIG. 123. The cloud chamber of C. T. R. Wilson (schematic).

electrodes by the field. The resulting potential pulses, after appropriate amplification by a proportional amplifier, are then fed to an oscillograph for recording, or in some cases simply registered by a counting device. The sealed chamber containing the electrodes, the gas, and the electric field is called, after Greinacher, an *ionization chamber.* The technical improvements of the ionization chamber, or the development of chambers for special purposes (for example, the double ionization chamber) are of no direct interest for our discussion. A modern ionization chamber is able to register more than one hundred particles per second.

The most popular detector and counter for nuclear particles and nuclear radiation is the *Geiger-Müller counter,* used today in millions of units all over the world. It combines high sensitivity with sturdiness and simplicity. This device is essentially an ionization chamber with a cylindrical metal tube as negative and a thin wire as a central positive electrode. The housing of the tube is charged to several thousand volts with respect to the wire electrode. A nuclear particle entering the counter produces ions and electrons by ionization, just as in the ionization chamber. However, near the central electrode the electrons are accelerated by the strong field to such an extent that they produce secondary ionization in the gas-filled tube. The total charge, greatly amplified by this secondary ionization, then is conducted to ground through a high resistance. The voltage pulse in the resistance can finally be measured by means of an amplifier (see diagram, Fig. 124). If the counter potential is not too high, the amplification of the primary ionization is proportional to the number of ions produced by the nuclear particle. Thus the particle energy can be measured. To detect very weakly ionizing particles, the potential of the Geiger counter can be increased to such an extent that the primary electrons start a sort of point corona discharge which, because of the space charge set up by the electrons, is extinguished after less than a hundredth of a second. The resulting voltage pulses in this case do not permit any conclusions to be drawn about the type or energy of the primary particles, but only about the frequency with which the particles enter the chamber. It should be noted that with the Geiger counter it is not difficult to measure the passage of every single particle into or through the chamber.

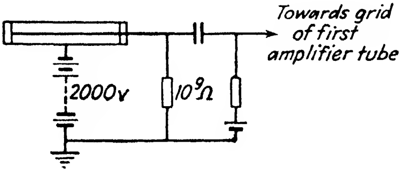


FIG. 124. Geiger-Müller counter with its circuit.

A *combination of a cloud chamber with one or more Geiger counters,* after Blackett, is very useful for studying the relatively rare cosmic-ray events. A counter can be connected in such a way that the cosmic-ray particle produces a voltage pulse which activates a relay. This causes the cloud chamber to expand, the lighting apparatus to function, and opens the shutter of the camera. A number of counters then can be so arranged ("coincidence circuits") that the nuclear particle which operates the cloud chamber must also pass through the chamber. In this way the cloud chamber can be automatically activated by the passage of a particle, or even by one entering from a definite direction. Arrangements of this and similar types are being used to a great extent with much success in cosmic-ray research (page 322).

Instead of Geiger-Müller counters, crystal counters are presently being used in increasing quantities. These counters are simply and ruggedly constructed and they require very little space. Such a counter consists of a crystal, to be penetrated by the radiation under investigation, with metal electrodes evaporated on two of its surfaces. The electrons released in the crystal by ionization migrate in the applied field to the anode and are recorded over an amplifier. The nature of the crystal and its minuteness are critical, since two conditions must be fulfilled. In the first place, the electrons released by an individual ionizing event must be able to migrate a sufficiently great distance before they are captured by lattice defects (see page 454), in order to give a sufficiently large current pulse. In the second place, the migration of the electrons must occur so rapidly that the counter after the shortest possible interval is ready for a new ionizing event. A crystal is also a convenient device for measuring the total ionizing radiation falling during a definite period on a particular location (for example, on a physicist). In certain ionic crystals (see page 458) absorbing "color centers" are produced by the radiation. The number of these centers, which can be measured optically, is proportional to the incident radiation.

Another detector of nuclear radiation, whose construction is based on the modern solid state research (Chap. 7), is the *scintilloscope*, which recently has been developed into an extremely sensitive device. It makes use of the fact that energetic particles can excite visible-light emission in certain solid substances. However, the minute single light flashes are not observed with a microscope (as was done 40 years ago with the scintillation method) but are amplified and recorded with a photomultiplier. With this type of apparatus 10^5 single flashes can be recorded per second and, furthermore, the kind of radiation and its energy can be determined from their intensity.

The *ionization produced by particles in a photographic emulsion* can also be used to detect and study nuclear processes, especially cosmic rays. This method was originated by Blau and Wambacher. A number of photographic plates are stacked together and, as in the cloud chamber, the energetic particles in passing through the stack ionize molecules, AgBr in this case, along their paths. When the plates are developed, tracks become visible and are studied under a microscope. With this method the kind and number of particles and in some cases even their energy can be determined. It is necessary to study the tracks of the particles with a microscope because the density of ionizable atoms in the photographic plate is about 1,000 times greater than in the gas of the normal cloud chamber. Thus each centimeter of a track in the cloud chamber corresponds to $\frac{1}{100}$ mm in a photographic plate. Furthermore, the time required for completely stopping a particle in the emulsion is about three orders of magnitude smaller than the time required to do the same in a gas. This is very important for certain cosmic-ray events, especially processes of mesons (page 331). The photographic method has been used more and more extensively in recent years, especially since the major photographic firms are producing special emulsions for this purpose, in fact special emulsions for the different ionizing powers of α-particles, electrons, etc. Figure 141 shows one of the first good photographs of this kind, taken in 1937 by Schopper.

5-3. Nuclear Projectiles and Their Acceleration

For the artificial excitation of nuclei, physicists use very fast, i.e., high-energy α-particles, protons, and deuterons (nuclei of the heavy hydrogen isotope, see page 45), in addition to neutrons and γ-rays. The positively charged α-particles, deuterons, and protons must have a very high kinetic energy in order to overcome the electrostatic Coulomb repulsion of the positive nucleus which is to be transformed. This repulsion does not exist for the uncharged neutrons, so that slow neutrons can penetrate into the nucleus and excite it or cause a nuclear transformation.

The projectiles used to transform nuclei (we can regard γ-ray photons as projectiles too) can be taken either directly or indirectly (see below) from natural radioactive sources, or they can be produced in discharges as helium or hydrogen ions and accelerated to the required high velocity by means of a high-voltage generator. Natural radioactive sources emit powerful α-rays of various energies (a maximum energy of 8.7 million ev is emitted by ThC′). The same sources emit energetic γ-rays (e.g., 2.6 Mev from ThC″), while γ-rays with the high energy of 17.5 Mev are obtained by bombarding the lithium isotope $_3\text{Li}^7$ with protons. However, protons and neutrons are not emitted from radioactive sources. Neutrons are produced, as we shall see on page 290, when beryllium $_4\text{Be}^9$ is bombarded with α-rays, so that a small tube with beryllium powder and an α-radioactive material is a convenient source for neutrons. For the production of artificial nuclear projectiles α, p, and d (p = protons; d = deuterons), these ions are pulled out of a powerful helium or hydrogen discharge by an electric field and then accelerated to the required velocity by a strong electric field. Energies as high as 400 Mev have been obtained in this way. This particle acceleration is accomplished today either in an acceleration tube (Fig. 125) with an applied potential difference of 1 to 15 million volts or in a machine in which the particles traverse the same small potential difference many times (multiple accelerators).

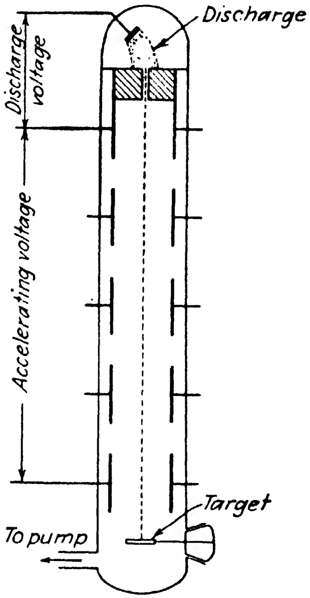


FIG. 125. Canal-ray tube with arrangement for further acceleration of the canal-ray particles (schematic).

To produce the high accelerating potentials of 1 to 15 million volts used in the first method, an *electrostatic Van de Graaff generator*, or a *cascade generator* charged by a transformer, is required. In the electrostatic generator designed by Van de Graaff (Fig. 126) a rapidly turning endless strip of insulating material is sprayed with a charge at C. The charge is then removed from the strip at F in the interior of the huge metal sphere A. A potential of only 20,000 volts can charge the metal sphere so highly that a corona discharge to the surrounding walls limits a further increase of the potential. In order to increase the useful potential without going to extreme dimensions for the sphere, the generator often is built into a pressure chamber which is filled with an electronegative gas (CCl_4 or CCl_2F_2) of several atmospheres pressure to reduce the loss through corona discharge. In the cascade generator (see Fig. 127) a multiplication of the potential furnished by the high-voltage transformer is effected by the proper series and parallel connection of condensers and rectifying tubes. Here again the available potential is limited by the size of the installation and the surrounding room. The cascade generator is, in general, more expensive than the Van de Graaff generator, which contains no expensive elements, but it does produce much larger currents and, consequently, more intensive particle beams.

Among the devices for producing multiply accelerated particles we dis-

tinguish between the linear and circular accelerators, depending upon whether the particles are accelerated in a straight line or are deflected by a magnetic field to follow a circular or spiral path.

The *linear accelerator* consists of a large number of cylindrical electrodes in a long tube (Fig. 128) which are connected alternately to a high-frequency potential source and to ground. The distances between the electrodes are so arranged that the particle under consideration is always in phase with the alternating field. Although the first linear accelerator was used before the first circular accelerator, Lawrence's famous cyclotron at Berkeley, interest in it soon waned because it required too long a tube and because it was very difficult to obtain suitable high-frequency power sources. However, during the war extensive experience in very efficient high-frequency generators such as those used in the development of radar was acquired. So, when after the war the desire for ever higher energy particles grew, interest in linear accelerators was revived. A number of these accelerators are now being built and tested at various places. We shall compare the advantages and disadvantages of the linear accelerator with respect to the circular accelerators after we have considered the latter in detail.

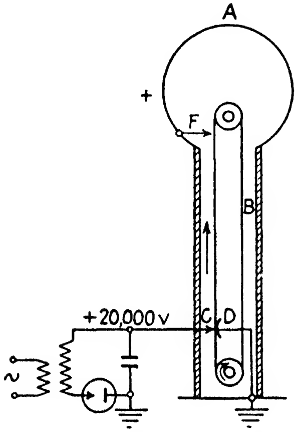


FIG. 126. Electrostatic Van de Graaff generator (schematic).

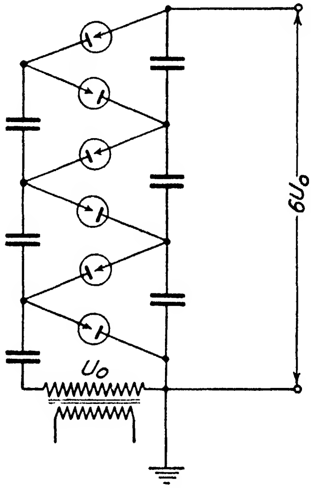


FIG. 127. Electric circuit of a cascade high-voltage generator (voltage multiplication).

The oldest circular accelerator is the *cyclotron*, developed by Lawrence. A flat metal can split in the middle, the halves called "D's" because of their shape (Figs. 129 and 130), is placed in a highly evacuated

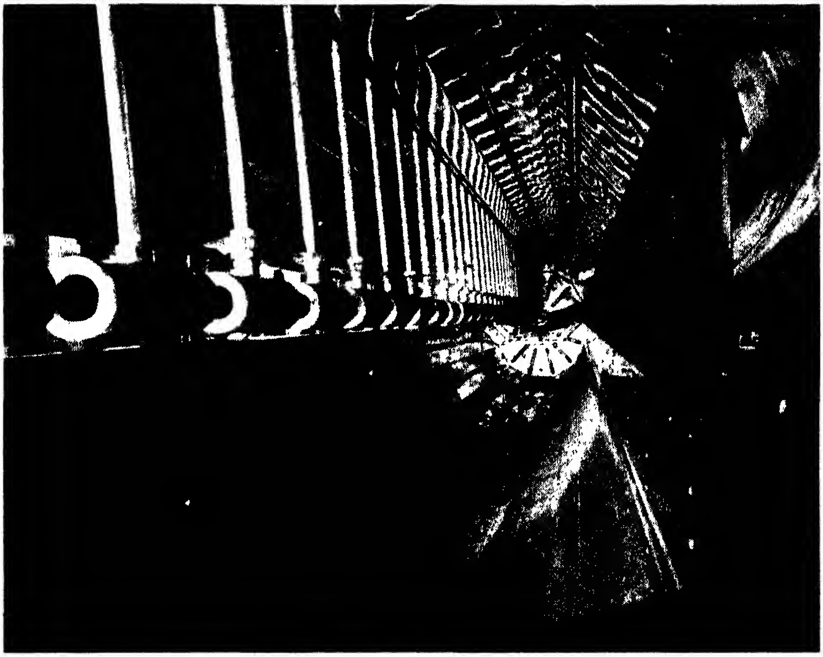


FIG. 128. Interior view of 40-ft acceleration chamber of the Berkeley linear accelerator. At the left are the drift tubes through which the beam of protons passes. Acceleration occurs as the particles pass the gaps between successive drift tubes. (*Photograph by courtesy of Radiation Laboratory, University of California, and the Atomic Energy Commission.*)

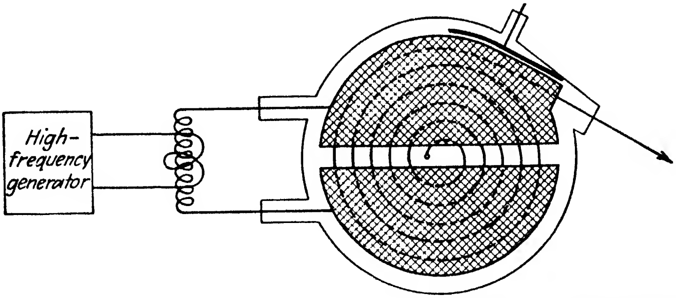


FIG. 129. Electric acceleration of ions in the D's of the cyclotron. The spiral orbit of an ion is shown.

chamber. The chamber is in a magnetic field of an electromagnet which is homogeneous up to some carefully calculated deviations. A picture of the largest cyclotron ever built is shown in Fig. 131. Its magnet alone weighs 4,500 tons. The two D's are connected to the terminals of a very efficient high-frequency power source so that the electric

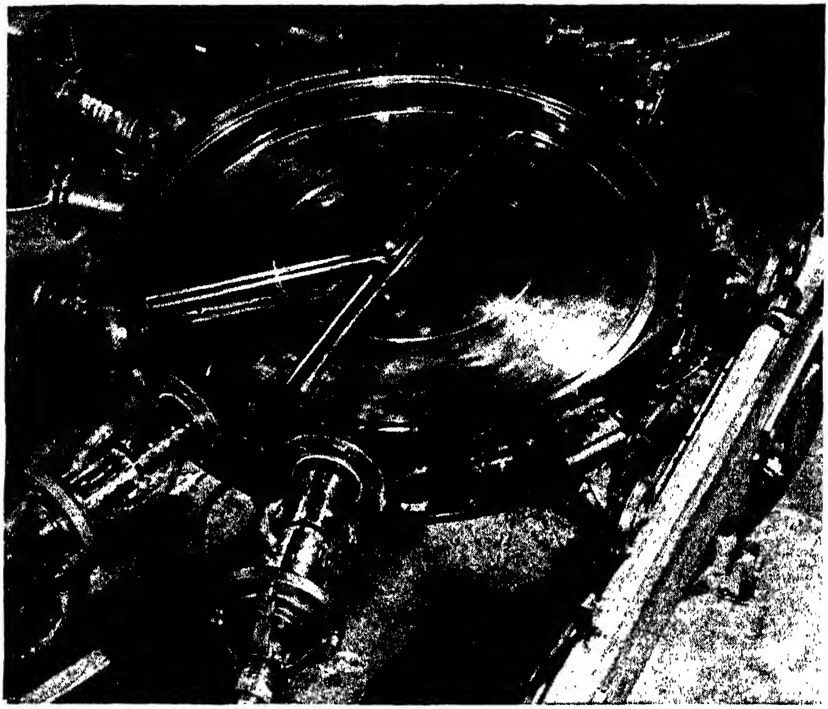


FIG. 130. Photograph of the D's of the cyclotron of Harvard University, 1939. (*Courtesy of Harvard University.*)

field in the space between the D's varies periodically with a high frequency (Fig. 129). The alternating voltage is about 10,000 volts. The ion source is placed somewhere between the D's. The ions of small initial velocity are now deflected by the vertical magnetic field into a circular path whose diameter is given by the equation we have frequently used

$$R = \frac{mcv}{eH} \tag{5-1}$$

which for a given magnetic field strength H depends only on the velocity v of the ions. If an ion in its path comes into the space between the D's, it will be accelerated by the electric field. The next semicircle of its path will then be traversed with a higher velocity, but, also according to (5-1), the path will have a larger radius, i.e., the length of the orbits and the velocity in the orbits will increase proportionally, so that the time elapsing between two traversals of the accelerating slit is independent of the acquired velocity; it is constant. The high-frequency source which produces the accelerating field and the magnetic

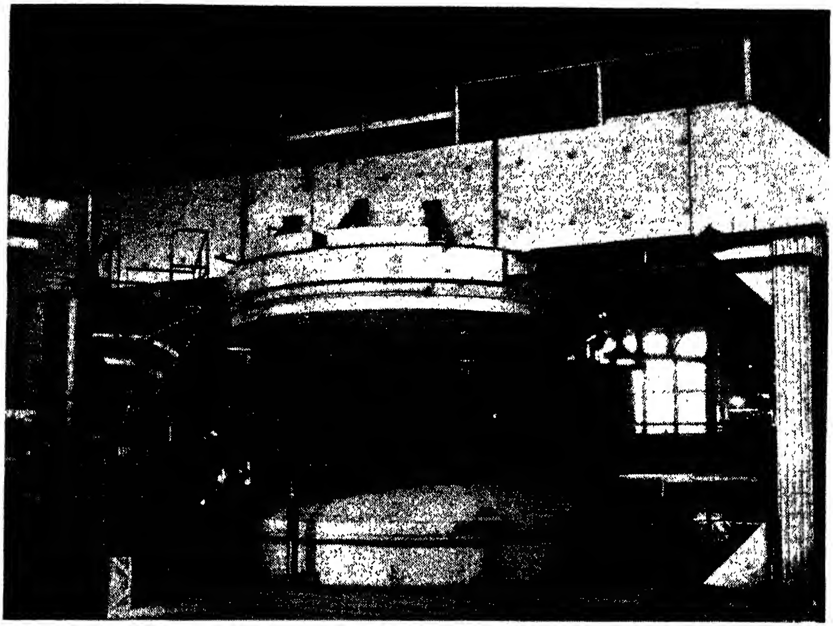


FIG. 131. A general view of the giant Berkeley cyclotron, now producing 380-Mev protons, 200-Mev deuterons, and 400-Mev α-particles operating as a synchro-cyclotron. Note the large radial crane of 30-ton capacity overhead, the 4,000-ton magnet with its 184-in. pole pieces, and the two 32-in. oil-diffusion pumps at left (only one visible) used for evacuating the rectangular steel vacuum chamber, which is seen between the pole pieces. This photograph was taken prior to the erection of an operating platform, and installation of 10-ft thick concrete radiation shield. (*Photograph by courtesy of Radiation Laboratory, University of California, and the Atomic Energy Commission.*)

field is so adjusted that the electric field alternates in phase with the rotational frequency of the ions. Then the ions will receive a 10,000 volt acceleration each time they pass through the slit between the D's and thus will describe a slowly opening spiral, until they arrive at the outer surface with an energy characteristic of the instrument and limited by the diameter $2R$ of the D's and the magnetic field strength H,

$$E_{\text{lim}} = \frac{m}{2}v^2 = \frac{e^2H^2R^2}{2mc^2} \tag{5-2}$$

The fully accelerated ions (protons or α-particles) can either strike a probe near the rim of the D's, or the ion beam can be deviated by a suitable magnetic field (deviating plate in Fig. 129) and made to pass through a slit into the outer space where it can be used for experiments. The cyclotron is very expensive because it requires a very heavy magnet. Most existing instruments have D diameters of 3 to 5 ft; the largest Berkeley cyclotron (Fig. 131) has a diameter of 184 in. Because it is necessary to tune the high-frequency source exactly to the magnetic field, the cyclotron is not an easy instrument to operate. However, it is one of the most efficient means of producing intense beams of high-energy particles. For example, with the 184-in. cyclotron, α-particles of 380 Mev, and recently, even protons up to 400 Mev have been produced. This is about 40 times the maximum particle energy which is emitted by natural radioactive sources! This largest instrument, however, is no longer an actual cyclotron. With its so-called *frequency modulation* it is closely related to the mammoth instruments, to be discussed below, with which still much more energetic nuclear particles can be produced. None of the latter instruments is actually in operation yet.

The cyclotron is not suited for the acceleration of electrons. The basic condition that the period of revolution is independent of the state of acceleration and that, therefore, the high-frequency field remains in phase with the accelerated particles requires, according to Eq. (5-1), a constant particle mass. This condition is sufficiently fulfilled for the heavy ions, and the small increase of the mass which follows from the relativity theory for high particle energies can be compensated by a slight field inhomogeneity in the outermost zones of the D's. The light electrons, on the other hand, at energies of only a few million volts acquire velocities which are nearly that of the light. The corresponding relativistic increase of the electron mass affects the electrons in a cyclotron so that they soon get out of phase; consequently the cyclotron is not suitable for the acceleration of electrons. There is, however, a considerable interest among physicists in the production of fast electrons. The many different ways of interaction of electrons with radiation and matter (X-ray production is one of them!) are by no means sufficiently known. So there is a need for experimenting with very fast electrons rather than making only accidental observations of them in cosmic rays (page 322). The extremely hard X-rays produced by them, furthermore, can induce nuclear transformations which also cannot be studied by other methods. Therefore, Steenbeck, starting from an old idea of Widerroe, developed an entirely new instrument for accelerating electrons, the betatron, with which electrons can be accelerated up to at least 100 Mev.

The fundamental principle of the *betatron* is simply that of the transformer (see Fig. 132). According to Maxwell's second equation

$$\frac{\mu}{c}\frac{\partial H}{\partial t} = -\text{curl } E \tag{5-3}$$

a varying magnetic flux induces a circular electric field, by which in the secondary of a transformer electrons are accelerated, i.e., an electric current is induced. In the case of the betatron the secondary coil is replaced by a high-vacuum tube surrounding the magnet core (Fig. 134). *Electrons which are injected into this doughnut-shaped high-vacuum tube are accelerated by the varying magnetic flux.* The principal difficulty in utilizing this simple idea is the problem of how to attain a stable circular orbit for the accelerated electrons which must encircle the core about 10^6 times without colliding with the walls of the tube in spite of the unavoidable collisions with the residual gas. The essential conditions for orbit stabilization were recognized by Steenbeck and taken into account in his construction. The varying magnetic flux is produced by an alternating current in an electromagnet built up of thin lamina. Between the ring-shaped pole pieces (Figs. 132 and 133) a magnetic *guiding field* is set up whose strength is always proportional to the actual (rapidly increasing) electron velocity. The function of this guiding field is to keep the electrons during acceleration in a circular path and to compensate for accidental deviations, arising from incorrect initial directions or from collisions with gas molecules, by forcing the electrons back to their required path. When the electrons have attained their maximum velocity (depending on the size and construction of the instrument), the guiding field no longer holds them in their circular orbit, so that the fully accelerated electrons leave their orbits and can be used for experiments, especially for exciting X-rays. Since the direction of the accelerating flux changes 500 times per second along with the alternating current which excites the field, the total

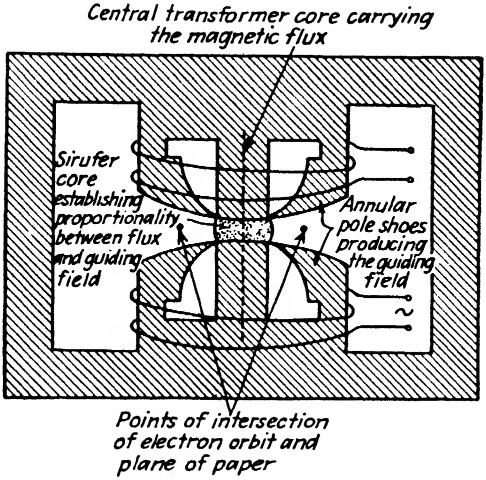


FIG. 132. Principal arrangement of the betatron, after the original patent of M. Steenbeck of 1936 (with minor changes).

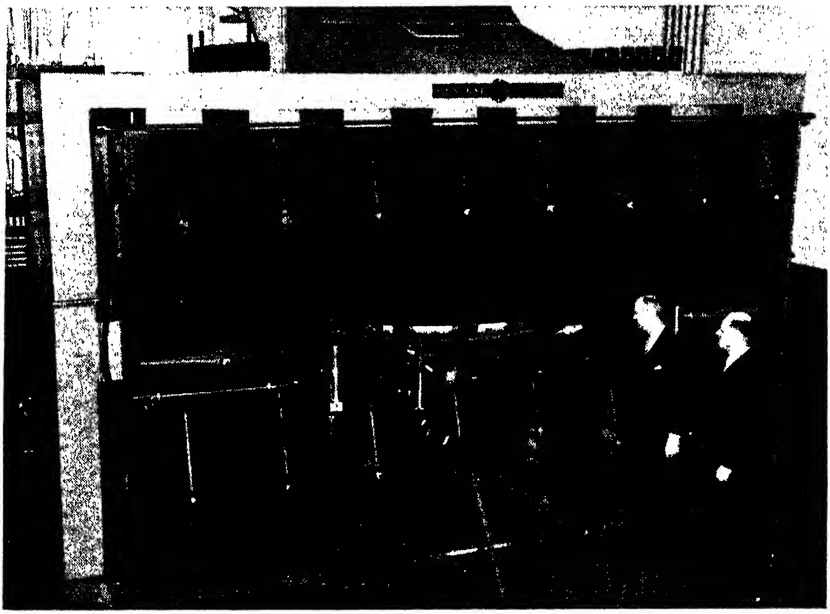


FIG. 133. Photograph of the 100-Mev betatron of the General Electric Company at Schenectady. (*Courtesy of General Electric Company.*)

acceleration of each injected electron group must be completed within one half of the period of the alternating current. Only if an electron group starts at the right instant can the maximum acceleration be attained. *The betatron thus does not produce a* continuous *stream of fast electrons, but each second a number of electron groups equal to the frequency of the accelerating field are accelerated.* We cannot go into a detailed description of the construction of the betatron. The first operating example of a betatron was built by Kerst. Small instruments only 2 ft long, which produce 6 Mev electrons and correspondingly hard X-rays, are already widely used for medical and technical purposes. Figure 133 shows the 100-Mev betatron of the General Electric Company.

A much higher acceleration of electrons in the betatron is limited by the fact that the electrons, in revolving in their cyclic orbit, radiate, to an increasing extent, energy of the frequency of revolution and its higher harmonics. The wavelength of the most intensive radiation is given as a function of the orbital radius R and the electron energy by the formula

$$\lambda_{\max} = \frac{2\pi R}{(E/m_e c^2)^3} \tag{5-4}$$

The electron acceleration by a betatron therefore becomes more ineffective with increasing electron energies and finally reaches a limit which according to theory is at about 500 Mev.

A further very effective acceleration is possible, however, if *the principle of the betatron is combined with that of the cyclotron.* In this method the electrons, in addition to being accelerated by the varying magnetic field, are accelerated by an alternating electric field between two electrodes as in the cyclotron. The *electron synchrotron* based on this idea thus combines in one apparatus the principles of construction of the betatron and the cyclotron: *As in the betatron the electrons are held in an orbit of constant radius R (in contrast to the cyclotron) by the guiding field and obtain their initial acceleration by the variation of the magnetic flux (inductive acceleration), while a further multiple acceleration to highest energies is produced by a cyclotron-like alternating electric field.* In order to hold the electrons in their orbits in spite of their increasing kinetic energy and to accelerate them at just the right instant, the change of the guiding magnetic field must occur synchronously with the acceleration of each injected electron group. For this reason this device, originated by McMillan and Veksler, is called the synchrotron. Figure 134 is a photograph of the doughnut-shaped vacuum tube (with the accelerating electrodes at the lower right) of a 70-Mev electron synchrotron of the General Electric Company. Except for the electrodes, the tube is the same as that of the 100-Mev betatron of Fig. 133. Synchrotrons have been built which accelerate electrons up to 300 Mev.

It is obvious that heavy particles (protons, α-particles) can also be accelerated with a synchrotron, if the initial inductive acceleration, which does not work too well with heavy particles, is replaced by injecting the particles into the synchrotron with sufficient initial energy (4 Mev) from a Van de Graaff generator. For large accelerators the synchrotron principle, because of its constant orbit, has the advantage that only an annular magnet is required. This reduces the necessary amount of iron, compared with a cyclotron of the same size, by orders of magnitude. *The largest such instrument under construction is designed to accelerate protons to* 10,000 *Mev* (10 *Bev* = 10^{10} *ev*). Figure 135 shows a sketch of this gigantic ion race track, which, because of its range into billions of electron volts and because it is hoped that artificial cosmic rays will be produced with it, is also called *bevatron* or *cosmotron.* The radius of the Berkeley bevatron will be 50 ft. It will take about 1 sec to accelerate a proton group from its initial 40 Mev to the final energy of 10 Bev, at which time the protons will have 99.6 per cent of the velocity of light and thus a relativistic mass which is more than 100 times their rest mass. Every 2 sec a new proton group will be injected. The task of adjusting the huge instrument so that the strength of the guiding magnetic field and the frequency of the accelerating electric field change synchronously with the acceleration of the protons seems a tremendous one. According to the plan, the frequency of the accelerating potential of only 12 kv will vary between 3.3 and 10 megacycles per second. There are 13,000 tons of iron and 400 tons of copper scheduled for the magnet, and 60,000 kva will be required to operate this huge proton race track.

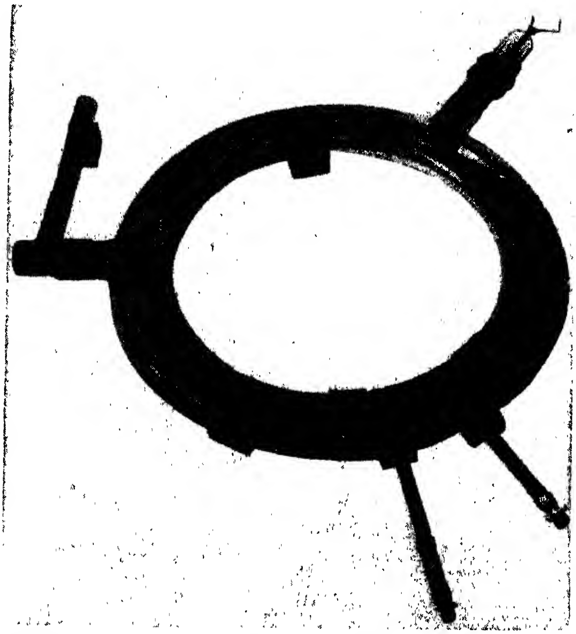


FIG. 134. Doughnut-shaped vacuum tube of the 70-Mev electron synchrotron of the General Electric Company at Schenectady. A similar tube is used in the betatron. (*Courtesy of General Electric Company.*)

We conclude this section with a short comparison of the advantages and disadvantages of linear and circular accelerators. For particle energies up to several hundred million electron volts the circular accelerator has the advantage of compact construction and a simpler high-frequency source compared to the very long linear accelerator which is difficult to adjust and requires a large number of high-frequency generators. The advantages of the latter, on the other hand, are the dispensing with the expensive magnet, the production of an automatically well-centered particle beam, and the lack of the unavoidable radiation of the charged particles in every circular accelerator which were discussed above for the betatron. Consequently, according to Slater, it is not impossible that linear accelerators for highest particle energies will replace or, at least, find a place in addition to the various modifications of the synchrotron.

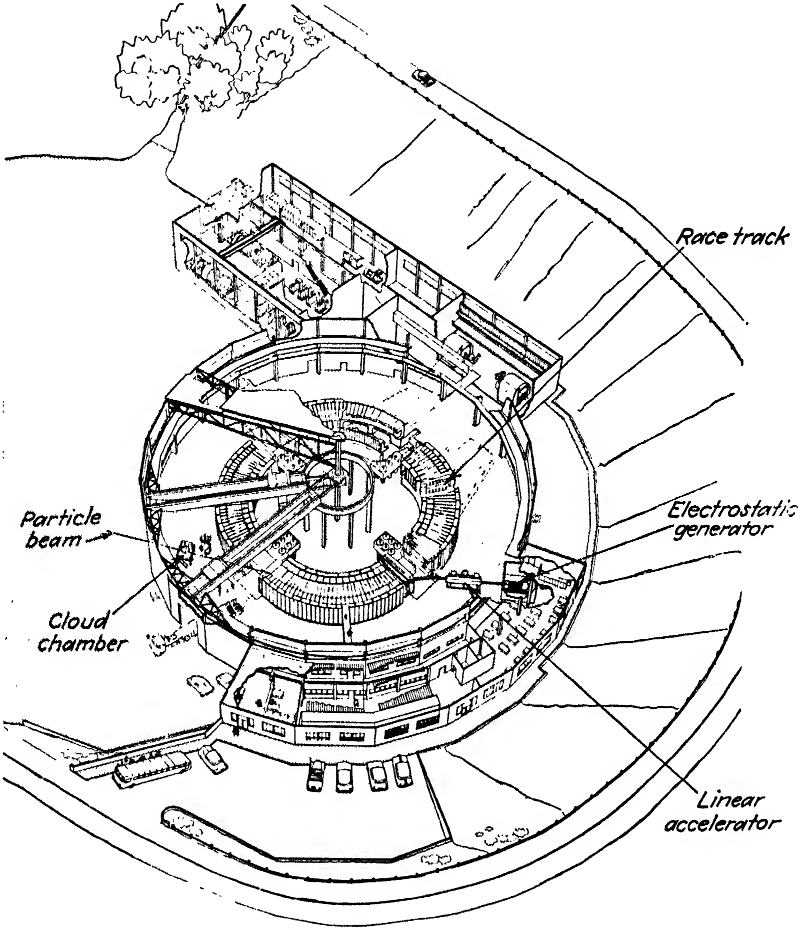


FIG. 135. Sketch of the proposed 10-Bev race-track synchrotron (bevatron) of the University of California. The ring-shaped magnet of this accelerator will be over 110 ft in diameter and will weigh more than 10,000 tons. As indicated in this sketch, according to the latest plan, the protons will be accelerated by an electrostatic generator to 4 Mev and then by a linear accelerator to 40 Mev before entering the race track. (*Courtesy of Radiation Laboratory, University of California, and the Atomic Energy Commission.*)

5-4. General Properties of Nuclei

Before we begin to discuss the variety of nuclear processes and their explanation we shall review what is known at present about the general properties of nuclei.

The *positive charge of the nucleus* can be determined from X-ray spectra (Moseley's law, page 26) or from the scattering of α-particles by the nucleus (page 24). It is equal to the atomic number of the element in the periodic table and is measured in units of the elementary charge e. This number is written as a subscript on the left of the symbol of the element, e.g., ${}_3\mathrm{Li}$.

The *mass of the nuclei* can be determined by the methods of mass spectroscopy (page 46) and is usually measured, not in grams, but in mass units, which are $\frac{1}{16}$ of the mass of the most abundant isotope of oxygen, O^{16}. For the relation of this so-called *mass number* to the physical or chemical atomic weight see page 44. The mass number is, as we shall see, identical with the number of nucleons (protons and neutrons) in the nucleus. The mass number is written as a superscript on the right (occasionally also on the left) of the symbol of the element, e.g., ${}_8\mathrm{O}^{16}$. We also know from page 54 that most elements (which as such are determined by their nuclear charge) have a number of different nuclei which are distinguished by their different masses. These are called isotopic nuclei. The most important of these is the heavy hydrogen nucleus of mass 2 which is designated by the symbol d to distinguish it from the proton p. We have already treated the methods of separation and of investigation of isotopes on pages 46*ff*.

The *size or diameter of the nuclei* is not well defined since a nucleus cannot be regarded simply as a rigid sphere. Its outer limits are determined by the decrease of the nuclear forces (compare the analogous situation with regard to the atom, page 18). We can determine the radius of the nucleus from Rutherford's formula, page 24, by measuring the angular distribution of α-particles scattered by the nucleus, if *we define the radius of the nucleus as that distance from the center of the nucleus at which the deviations of the nuclear forces from that given by the Coulomb repulsive force alone become noticeable* (*anomalous scattering*). Riezler, for example, determined, from such studies, the radius of the oxygen nucleus, O^{16}, to be 4.5×10^{-13} cm, and of the carbon nucleus, C^{12}, to be 4.1×10^{-13} cm. From these and similar measurements it follows that in general the radius of a nucleus of atomic weight A is given with an accuracy of 10 per cent[1] by the formula

$$r = 1.4 \times 10^{-13}\sqrt[3]{A} \quad \text{cm} \tag{5-5}$$

The author has called attention to the fact that the constant in (5-5) is, within the accuracy of the measurements, equal to the Compton

---

[1] According to a recent communication by Present, the number before the root in (5-5) is not a constant but varies with A from 1.30×10^{-13} for light nuclei to 1.48×10^{-13} for the uranium nucleus.

wavelength of the proton, h/M_pc. It can thus be expressed by the fundamental constants h and c and the mass of the proton M_p (see page 342). Additional methods for measuring the radius of the nucleus depend on the scattering of fast neutrons by the nucleus, and on the α-emission of radioactive nuclei which will be considered on page 278. In the last method the distance between the centers of gravity of the nucleus and the α-particle is computed at which the α-particle must have been ejected by the nucleus in order to reach its actual, measured velocity. All these methods, as well as several other indirect ones, lead to formulas of the type (5-5), and they all give constants lying between 1.3 and 1.8×10^{-13} cm.

According to (5-5) the volume of the nucleus is proportional to the atomic weight; that is, *the densities of all nuclei are approximately the same and equal to* 10^{14} *gm cm*$^{-3}$. Compared to the density of ordinary matter this is inconceivably large. The nuclei with respect to their constant density are similar to drops of the same liquid, whose density is also constant and independent of their radii. We shall use this *analogy between nuclei and liquid droplets* frequently in the following discussions.

Several statements can also be made about the *shape of the nuclei.* In a first approximation, we regard them as spherical. For the case of the light nuclei, this assumption is in good agreement with experimental results. However, according to Schüler, observations on the displacement of hyperfine structure components (see page 184) of elements of high atomic weight (heavy nuclei) indicate that there are deviations from the sphere in the shape of the heavy nuclei. An ellipsoidal shape agrees best with the measurements, but the deviation from the spherical shape, in general, appears to be of the order of only 1 per cent. In recent years the methods of microwave spectroscopy (page 72) have increased the accuracy of such measurements greatly. This deviation from spherical symmetry can be a prolongation or a shortening in the direction of the spin axis. Numerically this asymmetry is expressed most simply by associating an *electric quadrupole moment* with the nucleus and assigning a positive sign to the prolongation in the direction of the spin axis, and a negative sign to the shortening of this axis. For the nuclei which have been investigated, the numerical values of the quadrupole moment lie between -0.5 and $+6.0 \times 10^{-24}$ cgs units. For the case of the very heaviest nuclei we have evidence that the shape is even more elliptical. According to Hahn and Strassmann, the heaviest nuclei show a certain tendency to decay into two approximately equal parts. This will be discussed in detail on page 299. Such decay is hard to understand for the case of spherical nuclei, but much easier for an elliptical shape. Wefelmeier has developed rather detailed ideas of a geometrical (quasi-crystalline) arrangement of the nuclear particles in heavy nuclei, and from these considerations the asymmetry of the heavy nuclei can be explained.

Another important property of the nucleus is its spin. Just as the electron, the nucleus also has, in general, a constant mechanical rotation about its own axis. Like all atomic angular momenta (page 143), it is measured in units of $h/2\pi$ and is equal to $Ih/2\pi$ where I is an integral or half-integral number between 0 and $\frac{9}{2}$. We have already mentioned on page 182 that the nuclear spin is responsible for the hyperfine structure of the atomic lines, and we shall see on page 295 that it also causes intensity changes in the rotation bands of symmetrical molecules made up of like atoms, as well as the existence of the two modifications of molecular hydrogen (ortho- and parahydrogen). These modifications thus differ only by the different orientation of the nuclear spins of the two hydrogen atoms, which are parallel for ortho- and antiparallel for parahydrogen. From the spectroscopic investigation of hyperfine structure and band spectra an important law has evolved. *Nuclei of even-numbered atomic weight have an integral spin, or (in most cases) no spin at all; nuclei with odd-numbered atomic weight, on the other hand, have half-integral spin.* The proton and neutron each have a spin $\frac{1}{2}(h/2\pi)$. The importance of the spin for nuclear physics is based on the fact that due to its existence the law of conservation of the angular momentum is applicable also to the nucleus. We shall consider a number of important conclusions drawn from this law.

As in the case of the electron, a magnetic moment μ resulting from the mechanical spin is also associated with the nucleus. Whereas the magnetic moment μ_e of the electron is (at least in first approximation) equal to a Bohr magneton (page 154)

$$\mu_e = \frac{e}{2mc}\frac{h}{2\pi} \tag{3-84}$$

where m is the mass of the electron, the nuclear moment μ_n is measured in units of the nuclear magneton, which is obtained by substituting the mass of the proton for that of the electron in the above formula. Consequently, the nuclear magneton is only 1/1,837 of the Bohr magneton. Only recently has the problem of measuring the magnetic moment of the nucleus been satisfactorily solved, and at present there are three different methods for measuring this important quantity. In the Rabi method, a beam of atoms passes successively through different homogeneous and inhomogeneous magnetic fields, and from the effects of these fields on the atoms, analogous to the Stern-Gerlach experiment, the effective magnetic moments of the nuclei can be determined. In the second method, originated by Purcell, the material to be investigated for the measurement of the proton moment, for instance water, is brought into an orientating constant homogeneous magnetic field, perpendicular to which an adjustable high-frequency magnetic field is set up. If, during the continuous variation, the frequency of this field coincides with the frequency of precession of the nuclear moments about the constant field (Larmor frequency $\nu_L = \mu H/Ih$), the resonance produces a measurable reaction on the high-frequency circuit (increased damping), from which μ can be determined. In the third method, developed by Bloch, a receiver coil is adjusted perpendicular to the two fields used in Purcell's method, and an electromotive force is induced in them if the Larmor frequency and that of the high-frequency magnetic field resonate. This induced emf serves as a means of determining the magnetic moment of the nucleus. Actually, by these methods, the magnetic moment μ is not measured directly, but is determined from the Larmor frequency, which contains its ratio to the mechanical spin of the nucleus, I, the so-called gyro-magnetic ratio μ/I. Unexpectedly, the magnetic moment determined in this way is not an integral multiple of the nuclear magneton, but the values of μ_I are spread out continuously between -1.9 and $+6.4$. In particular, the proton has a moment of $+2.79$ nuclear magnetons and the neutron -1.91 nuclear magnetons. We shall return on page 338 to the meaning of this surprising deviation from the theoretically expected relation (3-93) (which is confirmed for the electron) between the spin and the magnetic moment. A negative value of the magnetic moment means that its direction and that of the spin are opposite.

Among the general properties of the nuclei, finally, the phenomenon of isomerism has to be mentioned, though it does not occur too frequently. By isomeric nuclei we mean nuclei which have equal nuclear charges and masses but different energies and degrees of stability, i.e., mean life-times (we speak now not only of stable nuclei). Nuclear isomers consequently must differ in the arrangement of their constituent particles, just as in molecular physics molecules of the same composition but with different arrangement of atoms are usually called "isomers." A nuclear isomer differs from an excited nucleus in that it does not revert to a normal state by emitting γ-radiation. Thus, referring to the terminology of atomic states (page 139), they can be called metastable. The reason for the forbidden transitions, similar to that in the case of the noble-gas atoms (page 134), probably is that the two isomers have different nuclear spins I, which cannot change easily enough by emitting radiation (compare the corresponding behavior of the electron spin, page 147).

5-5. General Discussion of the Structure of the Nucleus, Mass Defect, and Binding Energy

Before the discovery of the neutron, it was believed that all nuclei were built up from protons and electrons (A protons and $A - Z$ electrons). This idea, however, led to insurmountable theoretical difficulties. First, according to the uncertainty principle (page 198), an electron confined to the small space of the nucleus must have an exceptionally large momentum and thus an energy up to 10^8 ev, which is not consistent with our general knowledge of the nucleus. Second, the magnetic moment of the nucleus is about 2,000 times smaller than that of the electron. This is difficult to understand if electrons contribute their magnetic moments to that of the nucleus. Third, the mechanical spin of the electron as well as the proton is $\frac{1}{2}(h/2\pi)$. A nucleus of odd atomic mass number and even charge, like ${}_8O^{15}$, then should consist of 15 protons and 7 electrons and, consequently, have an even spin, in contrast to the experimental result. Heisenberg showed immediately after the discovery of the neutron that all these difficulties disappear if we assume the nucleus to be made up of protons and neutrons instead of protons and electrons. This idea of the nucleus consisting of Z protons and $A - Z$ neutrons can be regarded as certain at the present time. Since the protons, because of their positive charge, repel each other, a special nuclear force must be assumed to explain the cohesion of the nucleus. This force must act between neutrons and protons as well as between two protons and two neutrons. We shall consider the nature of this force on page 335. As we shall see on page 337, protons and neutrons in the nucleus transform themselves from one to the other and, consequently, may be regarded as two different states of the same particle. For this reason, these two constituents of the nucleus frequently are called "nucleons" as distinguished from the lighter particles, mesons, electrons, and neutrinos, which we shall meet in discussing nuclear transformations.

Even without a more thorough knowledge of the nuclear forces we can plot the potential in the vicinity of the nucleus, since we know that outside the nucleus the potential due to the charge $+Ze$ falls off as $1/r$, while inside the nucleus, as a result of the binding forces, there must be a potential trough. This trough is separated from the Coulomb repulsion potential on the outside by a potential wall as shown in Fig. 120. Thus the potential trough corresponds to the stable nucleus and each particle entering or leaving the nucleus must either overcome or penetrate the potential wall. We shall discuss the details of this when we discuss the radioactive decay on page 278.

What can we find out about the binding energy of the Z protons and $A - Z$ neutrons in the nucleus? Here we can derive information in a most interesting way from the so-called *mass defects of the nuclei.* If the atomic weight of the nucleus (i.e., the usual atomic weight diminished by the mass of the electron shells, $0.00054Z$) is compared with the sum of the masses of the protons and neutrons ($M_p = 1.0076$; $M_N = 1.0089$) which make up the nucleus, we find that *the mass of the nucleus is always smaller than the sum of the masses of its constituents.* The difference is called the mass defect. For example, expressed in mass units, the mass of the 2 protons and 2 neutrons of the ${}_2\mathrm{He}^4$ nucleus is 4.0330 and yet the mass of the He nucleus is only 4.0029. Thus the mass defect of the helium nucleus is about 0.03 mass units. Soon after the discovery of this fact, physicists related the mass defect to the law of the equivalence of mass and energy

$$E = mc^2 \tag{5-6}$$

and interpreted the mass defect in the following way: When a nucleus is formed from protons and neutrons, the binding energy must become free, since the result is a stable nucleus in a lower energy state than that of the nucleons before its formation. The mass equivalent of this energy is the mass defect. The binding energy, corresponding to this mass defect, of the He nucleus is, from (5-6),

$$0.030 \text{ mass unit} = 27.9 \text{ Mev} \tag{5-7}$$

if the nuclear energy is measured in million electron volts (Mev). For converting mass units into energy we have the relation

$$1 \text{ mass unit} = 931 \text{ Mev} \tag{5-8}$$

or, roughly, a mass defect of 1/1,000 mass unit is equivalent to 1 Mev. The converted energy is, as the example of the He nucleus shows, of the order of magnitude of millions of electron volts and this is about 10^6 times as large as the energy involved in processes of the electron shells of atoms and molecules.

The fundamental significance of the relation between mass defect and binding energy is that we have here for the first time a quantitative proof of the equivalence equation $E = mc^2$, which has been confirmed by nuclear physics in the exact form as required by the relativity theory. It can easily be seen that the mass equivalent of the binding energy of the electron shells, being 10^6 times smaller, is below the limits of observation. We shall meet further confirmations of the equivalence formula on page 282.

Let us return now to the binding energy of the nucleus. If the mass defects or the corresponding binding energies of all known nuclei are plotted against the number of nucleons (mass number A), we get in a first approximation a straight line, Fig. 136. This means that *the binding energy of each nucleon is approximately the same and equal to about 8 Mev.* This important result is in agreement with our droplet model of the nucleus, mentioned on page 267, since in a liquid droplet the binding energy of each attached molecule is the same. Our curve, Fig. 136, has some small fluctuations near the origin. A more detailed investigation actually shows that *the α-particle* (He *nucleus*) *with a binding energy of* 28 *Mev is the most stable existing unit which consists of several nucleons. The fact that these α-particles can be emitted by certain nuclei* (*radioactive α-decay*) *is in good agreement with these facts and seems to indicate that in the nucleus two protons and two neutrons prefer to form an α-particle instead of remaining single.* On the other hand, deuterons are comparatively less stable; their mass is 2.014 with a mass defect of only 0.003 mass unit and a binding energy of only 10 per cent of that of the α-particle. Consequently, *the α-particle is the only complex unit which may be formed as such in the nucleus and therefore can be emitted as an entity in nuclear transformations.* We shall go into all details of this when we treat the theory of nuclear structure on page 316. We shall take up next the internal processes in nuclei and their meaning.

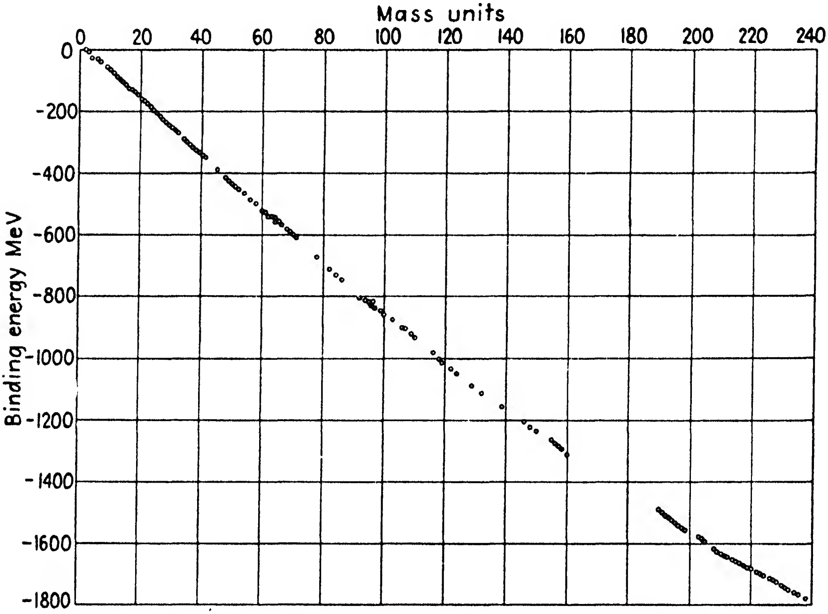


FIG. 136. The binding energy (in million electron volts) of all known stable nuclei plotted against their mass numbers.

5-6. Natural and Artificial Radioactivity and the Nuclear Processes Disclosed by Them

a. Natural Radioactive Decay

The first definite knowledge of the structure of the nucleus and its internal processes stems from the discovery of radioactivity in 1894 by Becquerel, and from the consequent radioactivity experiments of the Curies. It was realized soon that this phenomenon of spontaneously decaying nuclei of the heaviest existing atoms, which at first was so mysterious, is caused by an internal instability and that it occurs without any external cause and without any possibility of influencing it by normal physical means. We shall consider only briefly the empirical details, since these are discussed nowadays in every textbook of physics. In general, the product of the decay of a radioactive nucleus is another unstable nucleus, so that we find decay series whose end product is always a stable nucleus; in most cases it is a stable lead isotope. The three long-known radioactive series which begin with the uranium isotope 238, the uranium isotope 235, and the thorium isotope 232 and end with lead isotopes of mass number 206, 207, and 208, respectively, are shown in Fig. 137, together with a fourth series which has been found only recently. The mass numbers and atomic numbers of the sequences of nuclei which follow one another by the radioactive decay may be taken from Fig. 137. In addition to the kind of decay each of these nuclei is characterized by its half-life, which is inversely proportional to the probability of its decay. The half-life is the time in which half of the presently existing nuclei will have decayed.

Of greatest fundamental significance is the fact that the number of nuclei of one kind decaying per second depends only on the number of nuclei which have not yet decayed and is proportional to it,

$$\frac{dN}{dt} = -\lambda N \tag{5-9}$$

$$N = N_0 e^{-\lambda t} \tag{5-10}$$

Here λ is the so-called decay constant and is inversely proportional to the half-life of the particular nucleus. Thus this exponential decay law means that *the probability of decay of a nucleus does not depend on its age, as does the probability of the death of any living organism. The occurrence of such a decay process appears to be purely a probability matter.* We have discussed this already in relation to the theory of causality on page 244.

Three types of radioactive decay are known, according to whether

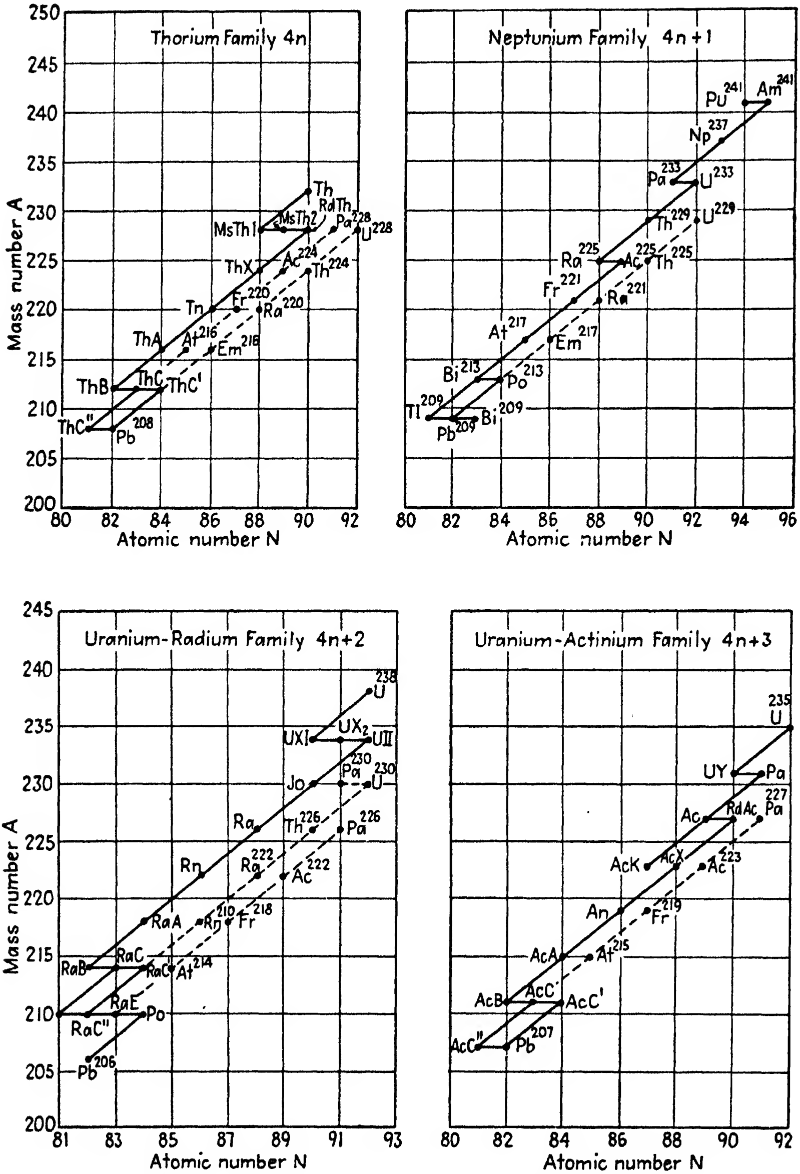


FIG. 137. The different radioactive families with their presently known side branches.

there is emitted from the nucleus a doubly positive helium nucleus (α-particle), a fast electron (β-particle), or a very energetic photon (γ-quantum). Each radioactive nucleus thus is characterized by its type of activity (α, β, γ), and its half-life or decay probability. From these facts, and from our knowledge of the structure of the periodic table, the displacement law of Fajans and Soddy, though quite a discovery at its time, is self-evident: *The nuclear charge (i.e., the atomic number) of a decaying nucleus is reduced two units by α-decay, leading to an element two places to the left in the periodic table. On the other hand, the positive nuclear charge is increased one unit by β-decay, so that the resulting nucleus is an isotope of the element on the right of the decaying element. Since γ-radiation does not imply a change in the charge of the nucleus, it is not associated with a transformation of an element.* This displacement law is evident from our illustration of the decay series. A line sloping to the left in Fig. 137 always corresponds to an α-transformation, whereas a horizontal line toward the right corresponds to a β-transformation (because it occurs without change of mass number).

One can see at once, from Fig. 137, to which elements the different radioactive transformation elements belong. Before their relation to the known elements became clear, each of these radioactive decay nuclei had been given its own name, which today can give rise to some misunderstanding. For example, we see from Fig. 137 that in the three old decay series there are seven different polonium isotopes with mass numbers 210, 211, 212, 214, 215, 216, and 218. Each of these has a different name, for example the istopes 210, 214, and 218 are called radium F, radium C′, and radium A, respectively, though they are isotopes of polonium and not of radium. A similar situation holds for the other decay nuclei, where, for example, mesothorium I and thorium X are radium isotopes, while mesothorium II is an actinium isotope.

In addition to the three decay series which have been known for a half century, some new decay series have been added as a result of recent research on artificially produced nuclei. These are the three series shown by the broken lines, all of which begin with a protactinium isotope and, after four or five steps, rejoin the original series, and the entirely new radioactive series which begins with the formerly unknown plutonium isotope Pu^{241}. It is shown in Fig. 137 as the second series from the left. Since the element of this series with the longest life-time is the neptunium isotope ${}_{93}Np^{237}$ with a mean life of 2.3×10^6 years, this new radioactive series is usually called the neptunium series.

It is of interest in connection with the special role which the α-particle with its four nucleons appears to play in the structure of the nucleus, that we now have exactly four radioactive series, whose members can all be represented by the following formulas:

$$\left.\begin{array}{lll}\text{Thorium series:} & 4n & (n = 58 \rightarrow 52) \\ \text{Neptunium series:} & 4n + 1 & (n = 60 \rightarrow 52) \\ \text{Uranium-radium series:} & 4n + 2 & (n = 59 \rightarrow 51) \\ \text{Uranium-actinium series:} & 4n + 3 & (n = 58 \rightarrow 51)\end{array}\right\} \quad (5\text{-}11)$$

b. The Explanation of γ-radiation

We begin the explanation of the radioactive processes by discussing γ-activity. Since γ-radiation is a shortwave electromagnetic radiation, it seems reasonable to draw an analogy between it and the processes leading to the emission of radiation by an atomic-shell electron, as discussed in Chap. 3. We thus assume that γ-radiation is emitted as a consequence of a change of energy of the nucleus, i.e., upon a transition of the nucleus from an excited state to the ground state, or in general to a lower energy state. Today, there is no doubt that this explanation is correct, and nuclear research is now directed toward determining the energy level diagrams for the nuclei corresponding to those of the atomic electron shells. *The mean life of an excited nucleus is much shorter than that of an excited atom because of the much greater interaction of the nucleons compared to that between the shell electrons and the nuclear charge. It is between* 10^{-13} *and* 10^{-17} *sec compared with the mean life of* 10^{-8} *sec of an excited atom.* Because of the much greater energies which are converted as a result of an energy transition in a nucleus compared with those of electrons, the γ-radiation has a much shorter wavelength. Since wavelength measurements of such short waves, especially of small intensity, are very difficult, and since techniques have only recently been developed to high accuracy, the term diagrams of the nuclei have not yet the same degree of accuracy which they have for atoms. These measurements can be carried out directly only for not too short wave and at the same time fairly intense γ-radiation by the methods of X-ray spectroscopy (page 73), for instance with the beautiful bent-crystal spectrograph of DuMond. In all other cases the wavelength determination is made by measuring the *energy* of a γ-quantum and then using the formula

$$E = h\nu = \frac{hc}{\lambda} \qquad (5\text{-}12)$$

to compute the wavelength. The energy measurements are made either coarsely by absorption in layers of material of different thicknesses, or, more accurately, by letting the γ-quantum be absorbed by an atomic shell and determining the kinetic energy of the resulting photoelectron which, except for the negligible ionization energy, is equal to the energy of the absorbed γ-quantum. The energy of the photoelectron then is measured by the *magnetic electron spectrograph* which is essentially identical with the mass spectrograph (page 46), except that now e and m are known and the velocity v is unknown. A check on and extension of the energy level diagrams of the nuclei determined in this manner is possible in the following way: An excited α-active nucleus, originating from a radioactive decay, can either (normal case) emit a γ-quantum, thus revert to its normal energy state, and then undergo a further α-decay, i.e., emit an α-particle. It can also (but seldom) dispose of these two transformations in one process by emitting a correspondingly more energetic α-particle (Fig. 138). In this case, in agreement with expectation, the energy difference of the two α-radiations is equal to the energy of the γ-radiation, i.e., equal to the excitation energy of the resulting nucleus. We shall return to further methods of extension and a more complete understanding of the term diagram on pages 292 and 321.

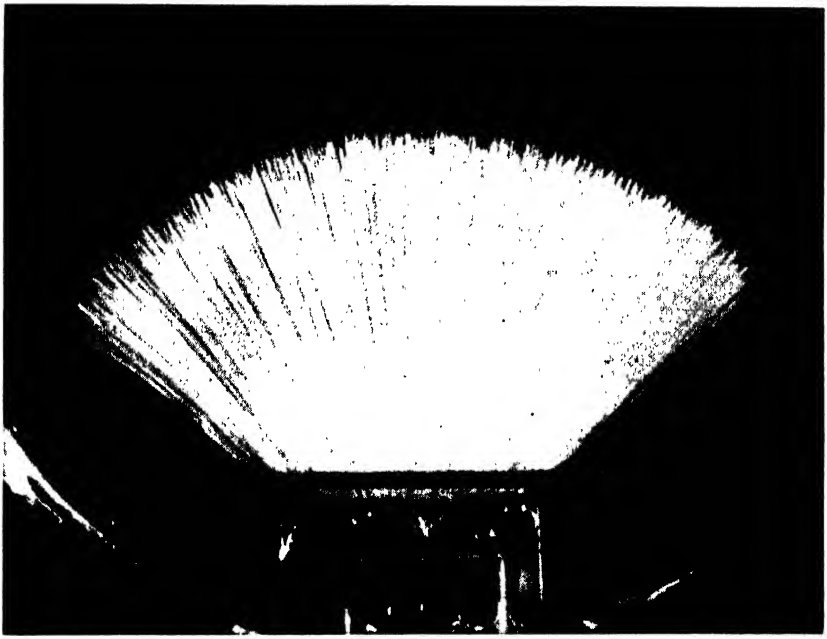


FIG. 138. Cloud-chamber photograph of a single high-energy α-particle of RaC′ among a large number of α-tracks of smaller (but constant) energy. The high-energy α-particle is emitted by an excited energy state of the same nucleus. (*After Philipp.*)

c. The Explanation of α-decay

We have already been concerned with α-decay, which we shall now explain. The energy of α-particles emitted by radioactive decay can be most easily determined by measuring their range in air. In passing through air the α-particles expend their energy by ionizing the gas molecules. The first clue leading to an understanding of the α-decay was the *Geiger-Nuttall relation.* It states that the range (energy) of α-particles is inversely proportional to the half-life τ of the emitting nucleus, or

$$\ln R = a + b \ln \frac{1}{\tau} \tag{5-13}$$

To explain α-decay we need the quantum-mechanical tunnel effect, discussed on page 234, which was discovered by Gamow in treating this case. As we have shown on page 270, the potential curve in and near the nucleus with reference to a positive particle such as the α-particle must be approximately that shown in Fig. 120, since a potential minimum corresponds to the stable nucleus. The potential outside the nucleus, at least from a distance of 3×10^{-12} cm on, must fall off as Ze^2/r because of the electrostatic Coulomb repulsion. The normal state of the nucleus corresponds to an arrangement of the nucleons in the lowest possible energy states in the potential minimum (analogous to the atomic shells, page 180). γ-radiation can be emitted by transition from an excited state of the nucleus to the ground state. Classically, an α-particle can leave the nucleus only if it has sufficient kinetic energy to surmount the enclosing potential wall. However, the Ψ-wave which, according to quantum mechanics, corresponds to the α-particle is not completely reflected, but can partially penetrate the potential wall, and the lower and narrower this is, the larger is its probability amplitude outside the nucleus. As soon as the α-particle has penetrated the potential wall, it is accelerated by the Coulomb repulsive forces; its measurable kinetic energy thus is the greater, the smaller the distance from the center of the nucleus at which it leaves the potential wall. If we call this distance r, then the kinetic energy of a doubly positively charged particle leaving a $(Z - 2)$-fold charged rest nucleus is

$$E = \frac{2(Z - 2)e^2}{r} \tag{5-14}$$

The Geiger-Nuttall relation is a consequence of this theory of the α-decay. To understand this we compare two α-particles, one of which, by the tunnel effect, comes from the ground state of the nucleus (Fig. 139) and the other from an excited state (or from a nucleus with a lower potential wall). The potential wall for the α-particle in the ground state is much higher and broader than that for an excited state. The decay probability of the normal nucleus, therefore, is much smaller and thus its half-life greater than that of a nucleus in an excited state. Conversely, the acceleration of the α-particle emerging from the higher state begins at A, that of the α-particle from the lower state begins at a greater distance, B. *The energy of the α-particle originating from the higher state with the shorter half-life thus is greater than that of the α-particle from the lower state with longer half-life (life-time).* This is just what the Geiger-Nuttall relation (5-13) states. The radioactive α-decay is thus explained and this explanation is in full agreement with the empirical results, an excellent accomplishment of quantum mechanics. It has already been mentioned that the distance r from the center of the nucleus, at which the acceleration began, can be determined from the range of the α-particles. By this r an upper limit of the nuclear radius is given.

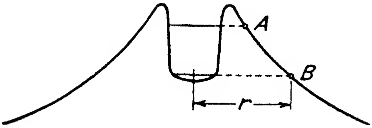


FIG. 139. The α-decay from two nuclear states (or two nuclei) of different energy. Used for explaining the Geiger-Nuttall relation.

According to Jentschke, a plot of the half-life values and the decay energies of all known α-emitters against their mass numbers reveals regularities which make it possible to interpolate and predict the half-life and decay energy of still unknown α-emitting nuclei.

d. *The Explanation of β-decay and the Existence of the Neutrino*

In our attempt to explain radioactive β-decay we encounter fundamental difficulties of two types. First, it is not directly comprehensible how electrons can come out of the nucleus (although we know there is no doubt about it), since the nucleus consists exclusively of protons and neutrons, or of α-particles made up of them. Secondly, the ejected electrons of the β-rays do not have, as do the α-particles (Fig. 138), a definite or, if there are several nuclear states, a few discrete energies. On the contrary, all measurements reveal a continuous velocity distribution of the β-particles, which is called the continuous energy spectrum of β-radiation. It increases to a maximum and then decreases down to zero energy. If we consider that the primary nucleus which is to be transformed by β-emission, just as every quantized atomic system, has a well-defined energy, and that the same holds for the final nucleus, then we are forced to the conclusion that the energy set free in the transformation must also have a definite amount. *Unless one assumes that in the case of β-decay the energy conservation law, which has been* *proved precisely valid for all phenomena of atomic physics, does* not *hold, one is forced to the only remaining conclusion that in β-transformation, in addition to the electron (β-particle), a second invisible particle is emitted.* In that case, the sum of its energy and that of the β-particle would be equal to the constant energy which is set free by the nuclear transformation. *This hypothetical particle which can have no rest mass or charge, because otherwise it could be observed in the cloud chamber, is called the* neutrino. The assumption of its existence removes another difficulty associated with β-decay. As we have already mentioned, nuclei of odd atomic weight have half-integral spin. Since the atomic mass number (number of nucleons) is not changed in a β-transformation, the spin must remain the same. On the other hand, the emitted electron has a spin of $\frac{1}{2}$, so that here we have another contradiction which is solved only if we ascribe a spin of $\frac{1}{2}(h/2\pi)$ to the neutrino which is supposed to be emitted simultaneously with the electron.

Thus the neutrino was first introduced as a working hypothesis with the only purpose of reestablishing the validity of the conservation laws for energy and spin, because these laws had proved correct in all other domains of atomic physics. However, physicists were not at all satisfied with this purely theoretical introduction of the neutrino, but looked for more evidence of its existence. Recently such evidence seems to have been produced. Namely, if a radioactive nucleus in a cloud chamber happens to emit a slow, low-energy electron, then according to the neutrino hypothesis a fast, i.e., high-energy, neutrino must be emitted. Its momentum must cause a recoil of the emitting nucleus which should be noticeable on the cloud chamber photograph, for it must be larger than if the electron alone were emitted. Investigation of recoils of β-emitting nuclei has been carried on in recent years and has produced positive results which are in agreement with our above expectation, so that *we can consider the existence of the neutrino to have been demonstrated.* From a computation of the total energy-mass balance of β-emission processes, to which we shall return later, we can determine the rest mass of the neutrino up to a certain limit. From these studies we are sure today that *the rest mass of the neutrino is less than $\frac{1}{300}$ of the rest mass of the electron, and probably is zero.* We shall have to deal more often with the neutrino, which has gained considerable importance in theoretical nuclear physics and especially in astrophysics, although occasionally its role is rather speculative.

Now we shall discuss the difficulty mentioned in the beginning, that in β-decay an electron leaves the nucleus although, to our knowledge, the nucleus does not "contain" electrons. Fermi tried to remove this difficulty by drawing a conclusion from the following analogy: In the case of the "transformation" of an excited atom into a normal atom a photon is emitted, although a photon as such did not previously exist in the atom. Similarly, according to Fermi, *an electron and a neutrino are emitted by the nucleus in β-decay although the electron (and the neutrino) did not previously exist as such in the nucleus. It is produced and emitted during the β-transformation just as a photon is "produced" and emitted during an electron transition in an atomic shell* (Chap. 3). By the Dirac theory, the emission and absorption of light (photons) by an atom is considered to be the result of an interaction of the atom with an electromagnetic high-frequency field. In a similar manner, as the photon corresponds to the change of the electromagnetic field, so, in the Fermi theory, the electron and neutrino are associated with the nuclear field in a manner which is not yet understood theoretically. The Fermi theory of β-decay met with certain quantitative difficulties (too small a life-time for β-active nuclei) when it was rigorously applied. These have recently been removed by an important refinement of the theory. According to this Yukawa theory, in β-decay not an electron and a neutrino are primarily emitted by the nucleus, but a meson, i.e., one of the "heavy" electrons found in cosmic rays (page 332), which has a mass 300 times greater than an electron and within 10^{-8} sec spontaneously decays in two steps into an electron and a neutrino. It seems that most β-decay experience is consistent with this refined theory, so that all phenomena of natural radioactivity are satisfactorily explained. On the other hand, there is lacking in the whole field of nuclear physics an exact theory of the nucleus by which the kind of decay which must occur and its probability can be derived quantitatively from the structure of a nucleus under consideration. Some theoretical approaches will be discussed on page 336.

e. Artificial Radioactivity. The Positron

Besides the natural there is an artificial radioactivity which was discovered in 1934 by Irene Curie (the daughter of the discoverers of natural radioactivity) and her husband, Joliot. They found that occasionally unstable nuclei are produced by externally induced nuclear transformations and then, as in the case of naturally β-active nuclei, revert to stable nuclei by β-emission. In contrast to natural radioactivity, where we have α- and β-transformations, the transformation of artificially radioactive nuclei occurs almost exclusively by β-emission, never by α-decay. Furthermore, as will be understood from nuclear theory discussed on page 320, *there are artificially radioactive nuclei which, in contrast to naturally radioactive ones, emit* positive *electrons, so that we have to distinguish between β^- and β^+ activity.*

These positive electrons, or positrons, which occur frequently in artificial radioactivity, were found by Anderson shortly before the discovery of artificial radioactivity in cloud chamber studies of cosmic rays. Except for the positive sign of the charge, the positron corresponds in all its properties to the negative electron. However, its mass may possibly be a little bit smaller than that of the negative electron. The obvious question, why only negative electrons had been found in nature, has a surprising answer. *The positron cannot exist long in the presence of matter. It unites with a negative electron, whereby the electron pair completely vanishes and, according to the equivalence relation* $E = m_0c^2$, *its kinetic energy and its mass are emitted in the form of two* γ*-quanta with a combined energy of* 10^6 *ev.* As we shall see on page 327, the reverse process is also possible and has been studied. A γ-quantum of more than 10^6 ev energy, upon interacting with the nuclear field, or in other words, by colliding with a nucleus, is transformed into an electron and a positron. This process is called *pair production.* Both processes are explained by Dirac's *hole theory,* which we shall discuss on page 328.

In addition to artificial radioactivity we should mention another form of nuclear transformation by which certain unstable nuclei can become stable and which sometimes competes with the β^+ decay. It is the so-called *K effect,* discovered in 1937 by Alvarez. An unstable nucleus with an excess positive charge (referred to the stable state of same mass number) can release this excess charge either by emitting a positron or can neutralize it by acquiring an electron from the electron shell. This effect actually occurs occasionally, e.g., with V^{48}, where it was discovered. This nucleus acquires an electron from the nearest electron shell, the *K* shell, and is transformed thereby into the Ti^{48} nucleus. *The "hole" left in the K shell of the* Ti *atom is filled by a transition of an electron from an outer shell. This transition* (*page* 119) *produces X-radiation of the K series, and these K lines are the means of detecting the K effect.* Meanwhile this effect has been found with a number of additional nuclei.

It may be mentioned that the same difficulties which were encountered in the theory of β-emission are also met in the *K* effect: Because the electron spin is half-integral, the law of conservation of total spin is fulfilled only if, simultaneously with the absorption of an electron, a second particle having a spin $\frac{1}{2}$ is either emitted or absorbed. It is highly improbable that a neutrino is absorbed since this would require that a neutrino be in the nearest neighborhood of the nucleus at the time of the electron capture. This requirement would make the *K* effect itself very improbable. Therefore, one is forced to assume that simultaneously with the absorption of the electron by the nucleus a neutrino is *emitted*. This neutrino, to distinguish it from that emitted with an electron, is called an *antineutrino*. This distinction is purely formal since the two seem to be identical. Also in the case of the K effect, the emission of an antineutrino has been confirmed by recoil studies of a nucleus which exhibits the K effect.

5-7. Artificially Induced Nuclear Transformations and Their Mechanism

Although naturally radioactive transformations occur spontaneously and cannot be influenced by physical means, transformations of normally stable nuclei can be *induced* by shooting high energy particles at these stable nuclei. As a result of experiments of this type our knowledge of the behavior of the nucleus has been greatly increased. The rapid progress of modern nuclear physics actually began with the discovery of these induced transformations in 1932. We arrive at a convenient theoretical model if we regard a nucleus as made up of a number of nucleons (protons and neutrons), enclosed by a potential wall (see Fig. 120, page 235). These nucleons are distributed, according to some rule as yet unknown, among the different energy states of the nucleus, as are the electrons in the atomic shells. If by collision with an accelerated particle (analogous to an electron collision with an atomic shell) or by absorption of a γ-photon such a nucleus acquires sufficient energy to overcome the binding energy of an individual nucleon in the nucleus, then, in contrast to the excitation of a shell electron, a number of nucleons will in general be excited to higher energy states, because the interaction of the closely packed nucleons is so much larger than that of the shell electrons. Because of this distribution of the excitation energy among a large number of nucleons we often speak of a heating of the nucleus. Just as by heating a liquid drop a molecule is occasionally evaporated, so, as a consequence of statistical fluctuations, one of the nucleons can acquire so much energy that it leaves the nucleus. This is analogous to the process of autoionization of the atomic shells as a consequence of double excitation (page 186). Such artificially induced transformations frequently lead to nuclei which are not stable but have more protons or more neutrons than is required by a stable nucleus (see Table 3, page 53). These unstable nuclei can revert to stable nuclei by emitting positive or negative electrons and neutrinos. This process of spontaneous transformation, which we discussed on page 281, is called artificial radioactivity. Artificially radioactive nuclei are, thus, such nuclei which after induced transformation have more or fewer protons than correspond to a stable nucleus.

However, like an excited atom, an excited nucleus can revert to the normal state by emitting radiation (photons), and this radiation by the nucleus, as we already know, is, because of the large binding energy of nucleons, high-energy γ-radiation. Thus we have arrived at the basic principles of all artificial nuclear transformations: If a nucleus is bombarded with high-energy projectiles (α-particles, protons, neutrons, deuterons, γ-quanta), they are either deflected (scattered) or absorbed. In the case of the latter there results, in general, either the emission of a nucleon, α-particle, or γ-quantum. The new nucleus resulting from the absorption of the colliding particle is called, after Bohr, the *intermediate nucleus.* The hypothesis of its existence as a physical reality makes it easy to understand that the colliding particle not only can knock out a particle (nucleon or α-particle) which already *exists* in the nucleus, but that, on occasion, a rearrangement of the nucleons in the nucleus occurs in which the bombarding particle is included. This applies, for example, to the bombardment of ${}_3\mathrm{Li}^7$ with protons, by which the intermediate nucleus ${}_4\mathrm{Be}^8$ is formed which, according to experiment, decays into two ${}_2\mathrm{He}^4$ nuclei, i.e., α-particles. In this case the proton shot into the nucleus must combine with two neutrons and one proton, already in the nucleus, to form the second α-particle, and the large binding energy that is set free by this α-formation causes the decay of the intermediate nucleus ${}_4\mathrm{Be}^8$ into two α-particles. *The nuclear reaction thus occurs in this, as in numerous other cases, in two steps, which may be regarded as direct evidence for the correctness of the assumption of an intermediate nucleus.*

The first induced transformation of this kind was discovered in 1919, though by no means intentionally, by Rutherford when he bombarded pure nitrogen with α-particles. A cloud chamber photograph of this transformation is reproduced in Fig. 140. The arrows point to the place where a collision between an α-particle and a nitrogen nucleus occurred. A long track originates here which, from the ionization strength, must have been due to a proton. Since the nitrogen contained no hydrogen, the proton must be the result of a nuclear transformation. The colliding α-particle, whose trace ends at the place where the collision occurred, must have penetrated into the nitrogen nucleus, and as a result of this collision a proton was ejected by the nucleus. Thus the charge of the nucleus is increased by one unit and the mass by three units: consequently, an oxygen nucleus ${}_8\mathrm{O}^{17}$, according to Table 3 the rarest of the stable oxygen isotopes, must have been produced from the nitrogen nucleus ${}_7\mathrm{N}^{14}$. We can write this nuclear transformation in the form of a chemical equation,

$$ {}_7\mathrm{N}^{14} + {}_2\mathrm{He}^4 \rightarrow {}_8\mathrm{O}^{17} + {}_1\mathrm{H}^1 \tag{5-15}$$

in which, because of the law of conservation of charge and mass, the

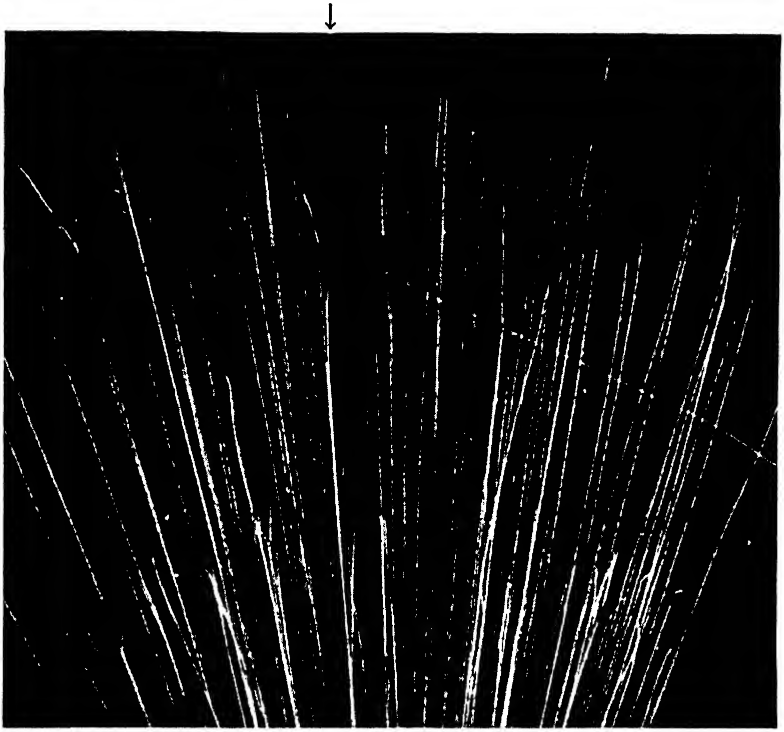


FIG. 140. Cloud-chamber photograph of the first artificial nuclear transformation, discovered by Rutherford. The emission of a proton (proceeding from the point of disintegration to the lower right corner) by a nitrogen nucleus after absorption of an α-particle. The nitrogen nucleus is transformed into an O^{17} nucleus which travels to the upper left due to the recoil. (*After Blackett and Lees.*)

sum of the units of charge (subscripts on the left) and that of the mass numbers (superscripts on the right) must be the same on both sides of the equation,

$$7 + 2 = 8 + 1 \text{ units of charge} \tag{5-16}$$

$$14 + 4 = 17 + 1 \text{ units of mass} \tag{5-17}$$

The momentum of the colliding particle is in this case shared by the ejected proton and the ${}_8O^{17}$ nucleus, whose trace is the short, thick line pointing to the upper left from the collision.

Following Bothe and Fleischmann, such transformations are written in a manner which is shorter and just as clear as (5-15) by indicating, in addition to the initial and final nucleus, the impinging particle and that ejected during the process. In this way we speak of an (α,p) process or an (n,γ) process. If the original nucleus is written before the parenthesis and the final nucleus after the parenthesis, the whole transformation is completely identified. Thus the transformation discussed above can be represented by the expression

$$_{7}N^{14}(\alpha,p)_{8}O^{17} \tag{5-18}$$

We mentioned above that the intermediate-nucleus hypothesis is necessary to explain many nuclear reactions just as the usual assumption of intermediate products in physical chemistry is used to explain the mechanism of a complicated chemical reaction. Just as in chemistry, the life-time of such an intermediate product may be so large that it can be measured. On the other hand, in certain cases it may simply be equal to the collision time (page 297).

The behavior of the nucleus discussed above describes correctly the reactions which are observed when nuclei are bombarded by projectiles of several million electron volts energy. Our pictorial concept also accounts for nuclear processes occurring in collisions with energies several orders of magnitudes higher, such as occur in cosmic rays. We shall discuss these on page 322. There are two limiting cases (though often difficult to distinguish) for collisions of 10^9 or 10^{10} Mev energy. In a central collision with a heavy nucleus (one with many nucleons) the colliding particle imparts its energy by numerous internal collisions to the nucleons, until its energy is expended and it is absorbed in the nucleus. This energy of 10^9 ev is equal to the binding energy of more than 100 nucleons. Its absorption causes, therefore, an intense "heating" of the nucleus, owing to which many nucleons are ejected. *In extreme cases the whole nucleus literally explodes.* Figure 141 shows an example of such a nuclear explosion in which the angular distribution of the nuclear fragments is isotropic, as has to be expected for an explosion of an overheated system. However, if the primary particle hits the nucleus well off-center, it interacts with only a few of its nucleons. In this case we cannot speak of an explosion, because there is no general heating, but of the ejection of some nucleons by the colliding particle. The particle directions then should show no isotropic distribution, but a definite preference for the direction of the colliding particle. This seems to be in agreement with observations. *For the very highest collision energy, finally, which is of the same order as the eigenenergy ($E = mc^2$) of the colliding particle, the concept of the intermediate nucleus fails completely, because now the binding energy of the nucleons in the nucleus is entirely negligible compared with the collision energy and the nucleus has to be regarded as a mere cluster of nucleons.* In this case we have to consider only collisions between the primary particle and the individual nucleons. The collision of a high-energy α-particle with a nucleus then has to be regarded as analogous to a simultaneous collision of four nucleons.

Let us return, after this excursion into what actually is the field of cosmic-ray processes, back to nuclear reactions due to collisions with particles of several million electron volts energy.

Each transformation of a stable nucleus is initiated or induced by a nuclear particle being shot into or attached to the nucleus to be transformed. As projectiles, only protons, neutrons, deuterons (i.e., the nuclei of the heavy hydrogen isotopes ${}_1H^2 = d$), and α-particles as well as high-energy photons (γ-quanta) play a considerable role. Of special importance is the possibility of inducing nuclear transformations by neutrons. Since these have no charge, they are not affected by repulsive electrostatic forces and so can penetrate even into the heaviest nuclei with the highest positive charge. We can show by a simple computation how much energy a charged particle must have in order to overcome the Coulomb force and penetrate into the nucleus, i.e., reach a distance of about 5×10^{-13} cm from the center of the nucleus. In order for a proton of charge $+e$ to approach through a distance r a nucleus of charge $+Ze$, an amount of energy

$$E = \frac{Ze^2}{r} \text{ ergs} \qquad (5\text{-}19)$$

has to be expended or, substituting the value of e and converting to million electron volts by the relation (3-10), 1 Mev = 1.60×10^{-6} erg,

$$E = (1.44 \times 10^{-13}) \frac{Z}{r} \text{ Mev} \qquad (5\text{-}20)$$

If we use the value 5×10^{-13} cm for

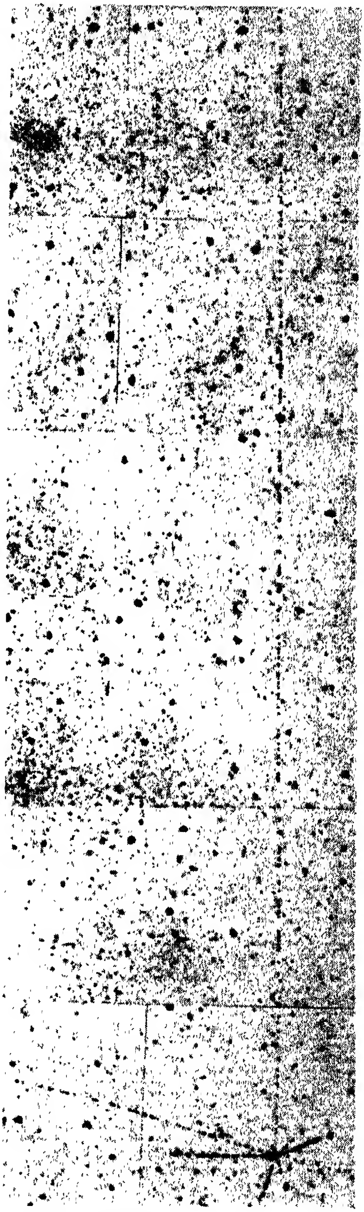


FIG. 141. Nuclear explosion, caused by a high-energy cosmic-ray particle. Photographed at 17,000 *m* altitude in 1937 by Schopper. ×250. (*Courtesy of E. Schopper.*)

r, then, to shoot a proton into a nucleus of atomic number $Z = 40$ (zirconium) against the repulsive forces alone, an amount of energy

$$E = 11.5 \text{ Mev} \tag{5-21}$$

must be expended. In order to shoot a proton into a zirconium nucleus it must be accelerated by a potential of more than 10 million volts. The superiority of neutrons, which are not affected by repulsive forces, for use as nuclear projectiles is evident from this computation. We mentioned already that nuclear transformations can also be produced by the absorption of high-energy photons (γ-quanta) by the nucleus. Such photons can be produced by the betatron or synchrotron (page 261) with energies up to 300 Mev with high intensity. Since in this process nucleons are released from the nucleus upon absorption of γ-radiation in the same manner in which electrons are released from their shells by the absorption of light (page 106), this process is called the *nuclear photoeffect.*

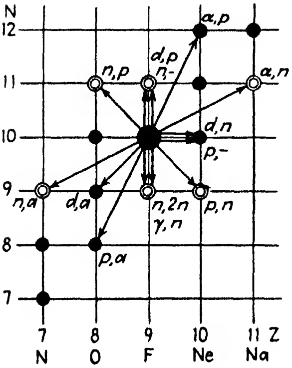


FIG. 142. Schematic representation of the transformations of the fluorine nucleus. Dots = stable nuclei in the environment of the fluorine nucleus, circles = unstable, radioactive nuclei. Near the transformation arrows the type of transformation is indicated by the abbreviated symbols. (*After Riezler.*)

No details are known about nuclear transformations induced by fast electrons. It is hard to separate them from the (γ,n) processes since sufficiently fast electrons, as we shall see on page 326, tend to produce high-energy photons (γ-quanta, bremsstrahlung). Consequently, we cannot distinguish whether the electrons or the γ-quanta produced by them are responsible for the transformations. We know, however, from cosmic-ray research that electrons of extremely high energy can cause nuclear explosions. The electrons, in this case, play only the role of energy carriers.

We see from our brief survey that there are for each nucleus a large number of possibilities for induced reactions. These can be presented by using a graph (Fig. 142), in which the number of protons in a nucleus (atomic number Z) is plotted as the abscissa and the number of neutrons $(A - Z)$ is plotted as the ordinate. Then each nucleus is represented by a point, with stable nuclei designated by black dots, unstable ones by small circles. An arrow is used to show to what nuclei the original nucleus can be transformed by reactions. In Fig. 142 all presently known artificial transformations of the fluorine nucleus, ${}_9F^{19}$, are indicated.

5-8. Production, Properties, and Detection of Neutrons

We have already referred, in the preceding sections of this chapter, to the neutron as the nuclear particle with no charge whose mass is approximately equal to that of the proton. But we did not go into the details of its discovery, its detection, or its production. We shall now take up this question after we have assembled the necessary information.

The existence of the neutron was inferred by Chadwick in 1932 in a bold, but physically logical, consideration from experiments of Bothe and Becker, 11 years after Rutherford had searched in vain for such an elementary particle. In 1930 Bothe bombarded ${}_4\text{Be}^9$ with α-particles and found a high-energy "radiation" which had surprising effects. This radiation was able to release protons of 5.7 Mev energy from substances containing hydrogen (see Fig. 143), whereas nitrogen nuclei of 1.2 Mev were released by it from substances containing nitrogen. The apparently obvious interpretation, that Bothe was dealing with very hard γ-radiation, and that the energy was transferred to the protons or nitrogen nuclei by Compton scattering (page 194), led to insurmountable contradictions. Computations using the laws of conservation of energy and momentum for scattering by protons and nitrogen nuclei gave two entirely different values for the energy of the γ-radiation, namely 55 and 90 Mev. Furthermore, the production of such high-energy γ-radiation by bombarding beryllium with α-particles could not be explained. Chadwick overcame all these contradictions by assuming the existence of an uncharged particle, which he called the neutron. By applying the laws of conservation of energy and momentum to a central collision of the hypothetical neutron with a proton and with a nitrogen nucleus, he obtained two equations for the mass and velocity of the neutron. Eliminating the velocity, he was able to compute the mass of the neutron. He found that, within the accuracy of Bothe's measurements, M_N was equal to the mass of the proton. A more exact determination of the mass defect by means of different nuclear reactions resulted in the following mass values (atomic weights) of the proton and neutron:

$$M_P = 1.007582 \pm 0.000004 \tag{5-22}$$

$$M_N = 1.008939 \pm 0.000007 \tag{5-23}$$

Since the atomic weight of the electron is $m_e/M(\text{O}^{16}/16) = 0.000548$, the mass of the neutron is, according to (5-22) and (5-23), greater than the sum of the masses of the proton and the electron. This means that the neutron can decay spontaneously into a proton and electron (+ a neutrino) and consequently must be radioactive. Its half-life is about half an hour, so that, because nearly all nuclei absorb neutrons readily and quickly, this instability of the neutron is not noticed in normal nuclear processes.

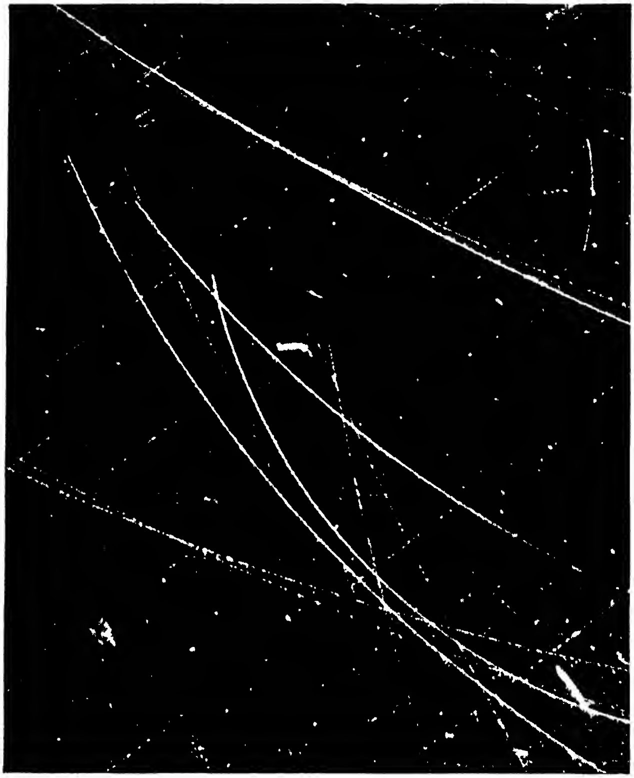


FIG. 143. Cloud-chamber photograph of protons, knocked out of hydrogen gas by impact of fast neutrons entering from the upper left corner. Proton tracks are curved due to magnetic field. (*Courtesy of Radiation Laboratory, University of California, and Atomic Energy Commission.*)

The Bothe transformation which led to the discovery of neutrons can be represented, in the terminology used on page 286, as

$$_4\mathrm{Be}^9\ (\alpha,n)\ _6\mathrm{C}^{12} \qquad (5\text{-}24)$$

This reaction is used today for producing neutrons when there is no high-voltage equipment available, and we have already mentioned on page 254 a neutron source consisting of a small tube containing a mixture of beryllium and an α-emitting compound. The energy, 13.7 Mev, of the neutrons emitted in this reaction is very high; unfortunately the scattering of the energy values is considerable. High-intensity neutrons can be produced by artificially accelerated particles. Special mention should be made of an artificially induced reaction which produces abundant neutrons by collision of deuterons of less than 0.5 Mev energy with deuterium compounds,

$$_1H^2\ (d,n)\ _2He^3 \tag{5-25}$$

In addition to this reaction, the bombardment of lithium, beryllium, boron, and carbon by fast deuterons produces neutrons by (d,n) transformations. Of special interest to nuclear research (see page 299) is the production of thermal neutrons, i.e., of neutrons whose kinetic energy is equal to their thermal energy kT at room temperature. Thermal neutrons are obtained when fast neutrons are stopped by substances containing hydrogen, such as water or paraffin, in which the neutrons give up their kinetic energy in collisions with protons. However, since protons have a tendency to absorb slow neutrons to form deuterons, $_1H^2$, the thickness of the retarding hydrogen-containing material must be so adjusted that the retardation is sufficient without resulting in the absorption of too many neutrons. If it is necessary to suppress completely any absorption of neutrons, heavy water, D_2O, or carbon must be used as the "moderating" substances because these materials, or rather their nuclei, are among the few known nuclei which do not absorb neutrons. As we shall learn on page 306, this method of stopping neutrons now is used widely for the production of atomic energy. The "piles," to be discussed later, are also the best sources of slow neutrons that we know.

Since the neutron, because its mass is about equal to that of the proton, can transfer its whole energy in colliding with a proton, and since slow neutrons, furthermore, are absorbed by protons, a neutron beam is completely absorbed by the protons of a layer of about 20 *cm of water or paraffin.* For this reason, a water shield is a very efficient protection against neutron radiation. On the other hand, according to the law of momentum, *neutrons impart only about* 2 *per cent of their energy to heavy nuclei such as lead, so that we have the apparent paradox of fast neutrons penetrating lead without any considerable reduction in their energy.* The contrasting behavior of lead and paraffin toward charged particles, on the one hand, and toward neutrons, on the other, may occasionally be of value as a method of separating beams of neutrons and protons.

Since the neutron has no charge, it does not ionize in colliding with atoms or molecules. Consequently neutrons cannot be detected directly in the cloud chamber, ionization chamber, or Geiger counter. To detect neutrons, and to measure their energy, the ionization by protons which are knocked out by neutrons from substances containing hydrogen is used. For neutron studies, therefore, the cloud chamber, ionization chamber, or Geiger counter is either filled with hydrogen or some hydrocarbon under high pressure (Fig. 143) or its walls are coated with paraffin. To detect very slow and, especially, thermal neutrons, which do not have sufficient energy to knock out protons, another process is used. Many nuclei absorb slow neutrons which then induce a transformation, especially the (n,γ) reaction, by which radioactive, β-active nuclei are produced. The β-rays emitted in their conversion then serve as a measure of the intensity of the neutrons. Cadmium and silver are especially sensitive neutron indicators of this type.

We have already mentioned, on page 287, the most important property of neutrons for nuclear chemistry: Since neutrons have no electric charge they have no repulsive forces, i.e., there is no potential wall between them and the nucleus they approach. Thus even the slowest neutrons can penetrate into *all* nuclei. Moreover, this process has a high probability because the interaction time for a thermal neutron and a nucleus is much greater than in the case of an impact of a fast particle.

5-9. The Energy Balance of Nuclear Processes and the Energy-level Diagrams of Nuclei

Like chemical reactions, a nuclear reaction is completely known if in addition to the reaction formula the energy balance associated with the reaction, the "heat of reaction," is known. Very closely connected to the problem of energy balance of nuclear processes is that of determining the possible energy states of the nucleus, i.e., the energy-level diagram, and the transition rules for the different states.

It is clear that *a nucleus, as a system consisting of a certain number of nucleons, is subjected to quantum rules in the same way as an electron shell, which is a system of a certain number of electrons.* Thus we can expect a nucleus to have a series of discrete energy levels corresponding to the stationary energy states of the nucleus, as well as one or more continuous energy regions corresponding to the decay processes of the nucleus, comparable to the ionization continuum of the atom (page 106) or the dissociation continua of a molecule (page 377). However, in contrast to the electron shells, the nucleus has no central force field. Each nucleon moves in the resulting field of all the other nucleons. *In contrast to the Rydberg series of the atom (Fig. 57, page 100), we do not expect the nucleus to have a regularly arranged system of energy states.* Because of the strong coupling between the nucleons, the excitation energy of the nucleus can distribute itself over many, in the limiting case, over *all* nucleons of the nucleus ("heating" of the nucleus, page 283), so that there are corresponding stationary energy states of the nucleus whose energies are far greater than the separation energy of an individual nucleon or α-particle. *Only in rare cases (ground states of stable isotopes) are nuclear energy states actually stable.* In general, an excited nuclear energy state can make a transition to a more stable state, often that of a *different* nucleus, not only by radiating (γ-emission) but also without radiation (see autoionization, page 186), e.g., by electron emission (β^+ or β^-), electron acquisition (K effect, page 282), emission of an α-particle, proton, or neutron. In determining the energy-level diagram of an atom we had to consider only that one atom. However, in determining the energy-level diagram of a nucleus, we must, as we shall soon see, consider the numerous possible transitions between the nucleus under consideration and its neighbors.

The excitation of a stable nucleus is possible either by absorption of a γ-quantum, by the acquisition of an electron (K effect), or by incorporating an impinging proton, neutron, or α-particle, since each incorporated nucleon makes available as excitation energy, in addition to its kinetic energy, an average binding energy of 8 Mev.

Let us consider more precisely this relation in the case of the (p,α) transformation of ${}_3\mathrm{Li}^7$. It is known from cloud chamber experiments that when a ${}_3\mathrm{Li}^7$ nucleus is bombarded with protons of 0.4 Mev kinetic energy, the intermediate nucleus produced, ${}_4\mathrm{Be}^8$, decays into two α-particles, each of which carries away 8.8 Mev kinetic energy. We can represent this process in the following manner

$${}_3\mathrm{Li}^7 + {}_1\mathrm{H}^1 + 0.4\ \mathrm{Mev} \rightarrow 2\,{}_2\mathrm{He}^4 + 17.6\ \mathrm{Mev} \qquad (5\text{-}26)$$

We can compute this reaction energy of $17.6 - 0.4 = 17.2$ Mev. It is in good agreement with the experimental results. Because of the equivalence law $E = mc^2$ *the masses on the left side of* (5-26), *plus the kinetic energy divided by* c^2, *must be equal to the sum of the masses on the right side, plus the kinetic energy of the reaction product divided by* c^2. That this is true can be seen by inspecting Eq. (5-26). In the chemical terminology this nuclear reaction is strongly exothermal; the large energy of 17.2 Mev or 398 million kcal/mole is released. Just as in chemistry, there are also endothermal nuclear reactions. In these reactions the kinetic energy carried into the nucleus by an injected particle is greater than the kinetic energy set free. Because of the complete analogy of these nuclear reaction equations to the corresponding reaction equations of chemistry, this branch of nuclear physics which deals with the determination, by quantitative experiments, of all possible nuclear reactions and the kinetic energy freed or trapped, is called *nuclear chemistry.*

Occasionally one finds, for a definite reaction, instead of *one* definite reaction energy a number of different values. The largest corresponds to the formation of the final nucleus in its normal state, whereas the smaller values correspond to the formation of the final nucleus in various excited states. This fact can be experimentally confirmed by the result that the energy of the simultaneously emitted γ-rays agrees with the difference of the reaction energies and thus with the difference of the energies of the different nuclear states.

We discuss the relation between the reaction energies and the energy levels of the participating nuclei in more detail by making use of a graphic representation by Pollard. In Fig. 144 the excited states of an initial nucleus and an approaching α-particle are shown. By the absorption of this α-particle a new nucleus (intermediate nucleus) is produced, with a different energy-level diagram, indicated in Fig. 144*c*. As a result of the absorption of the kinetic and binding energies of the α-particle, this intermediate nucleus is highly excited. In our example it is transformed, by emitting a proton, into an end nucleus which has its own term diagram. After emitting a proton, the final nucleus may be in the ground state or in any one of its excited states. The kinetic energy of the emitted proton must be different in these cases, and the same applies for the reaction energies Q_1, Q_2, Q_3, or Q_4, as plotted in Fig. 144. Their differences must agree with the term differences in Fig. 144*c*. Thus the observation of nuclear processes of different reaction energies makes it possible to determine certain energy states of the final nucleus. The differences of the states, according to Eq. (5-12), are observed as γ-radiation, if the corresponding transition is "permitted." Consequently, we know only those energy states of the nuclei

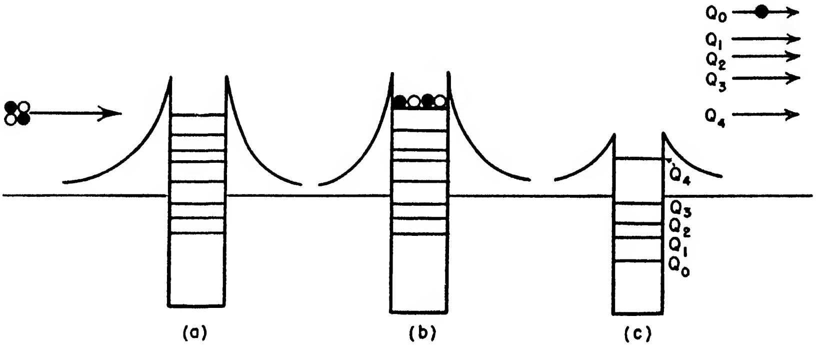


FIG. 144. Schematic representation of an (α,p) process. The α-particle approaches and enters the nucleus. By its inclusion a new (intermediate) nucleus with a different set of energy levels is formed, which then emits a proton. Depending on the energy level from which the proton is leaving, different Q values are found. (*Courtesy of E. C. Pollard.*)

which can be observed by their γ-radiation or by processes of the indicated (though often more complicated) type. Our knowledge of the possible nuclear energy states is very meager. For the nuclear processes which are accessible to us are limited not only by selection rules, but are also, for experimental reasons, restricted to certain energy regions. A survey of the term diagrams of several nuclei according to the latest data of Pollard is given in Fig. 145, while Fig. 146 shows the levels of the Be^8 nucleus with the processes used for their determination by Lauritsen. For example, a normal Be^8 nucleus and one which is excited to 3 Mev is produced by the (p,α) process of the B^{11} nucleus, indicated at the right side, with an expenditure of 0.82 Mev of kinetic energy. The yield curves shown in the figure will be of interest to us in the next section.

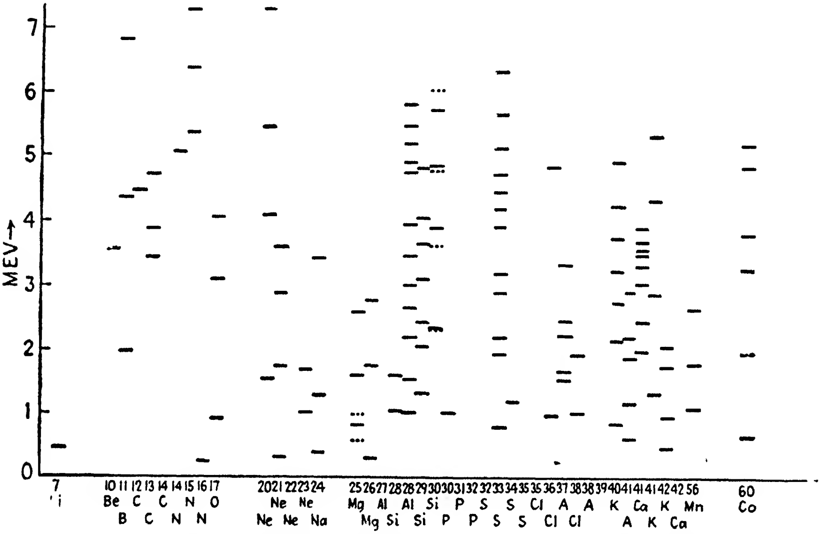


FIG. 145. The presently known energy states of a number of atomic nuclei with atomic numbers between 7 and 60. (*Courtesy of E. C. Pollard.*)

The *width of the nuclear energy levels* can be measured directly; it ranges, because of the different transition probabilities for nonradiative transitions and for radiation, between the wide limits of approximately 10^{-1} and 10^5 ev. The corresponding *life-time of nuclei in these states* can be computed from the uncertainty principle, Eq. (4-22). It turns out to be between 10^{-14} and 10^{-20} sec.

An interesting application of these considerations concerns the intermediate nucleus B resulting from the penetration of a colliding particle into the nucleus A. The energy of nucleus B depends on the kinetic energy of the absorbed particle and its freed binding energy. If the

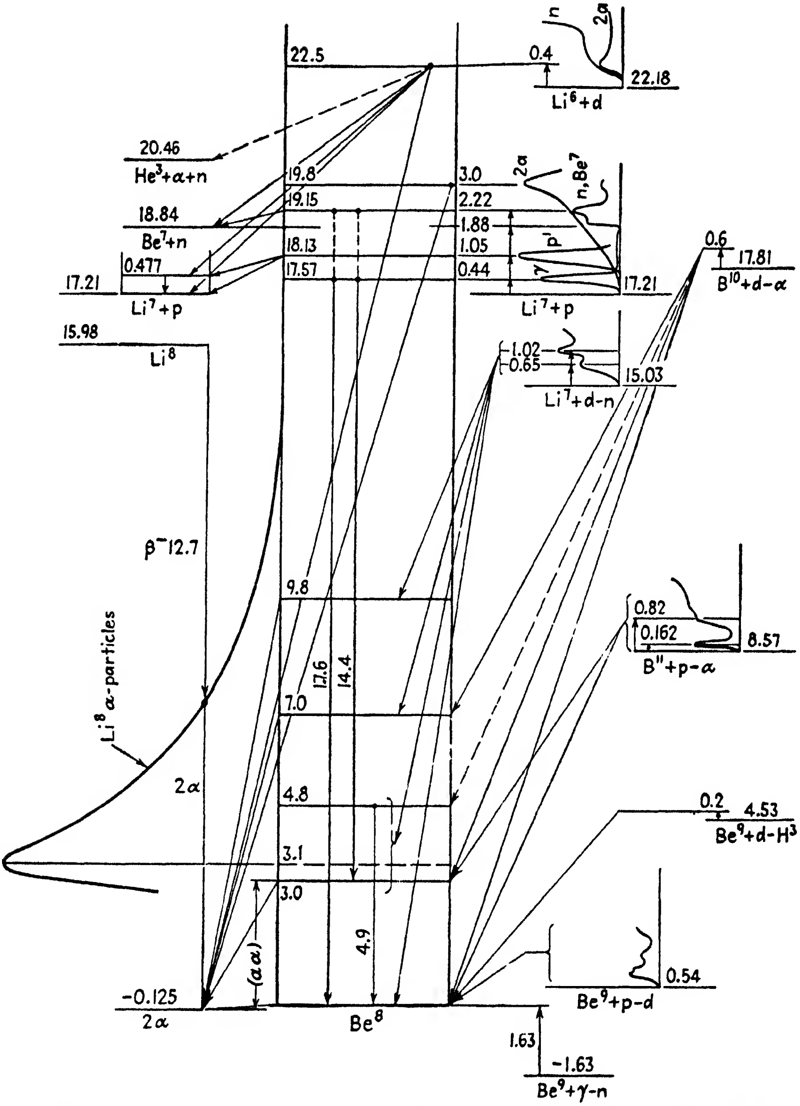


FIG. 146. The energy-level diagram of the $_4Be^8$ nucleus according to the presently known information. (*Courtesy of T. Lauritsen.*)

total energy of this intermediate nucleus corresponds to one of its quantized energy states, then it has a life-time corresponding to the width of this state. If it does not, then the colliding particle cannot be absorbed and the intermediate nucleus cannot exist as a physical reality. In this case we have only a collision between the particle and nucleus A. Its duration can be computed, from traversing the nucleus of diameter of 10^{-12} cm with the particle velocity of about 10^9 cm sec^{-1}, to be 10^{-21} sec.

Our brief discussion has made it clear that the most complete determination possible of the energy levels of all nuclei, including the pertinent selection rules for all transitions (as it has been in the case of the atoms and molecules), is one of the most important prerequisites for the quantitative theoretical description of the structure of the nuclei. As in the case of electron shells, this theory will lead to the characterization of the different nucleons and their states by quantum numbers. There are already indications for the existence of closed nucleon shells, corresponding to the closed electron shells of the noble-gas atoms (page 169), since there are *certain "magic" nucleon numbers which correspond to particularly stable nuclei.* We return to this briefly on page 321. This branch of nuclear physics thus seems especially suited for making clear the close interrelations of the different fields of atomic physics.

5-10. The Yield of Artificial Nuclear Transformations

After we have discussed the possible nuclear transformations under emission of individual particles or γ-photons, their mechanism, and the energy converted in these processes, there remains to be considered briefly the yield of these artificial transformations. The yield always depends on the nature and the velocity of the colliding particles, and it can be characterized by the number of particles which, having a given velocity, must be shot into a thick layer of a substance in order to produce one transformation. For those nuclear reactions which have the greatest yield, this number is of the order of 10,000; but for most reactions the number is 2 or 3 orders of magnitude higher.

Instead of this figure of the yield, it is preferred in atomic physics to characterize a collision process by the effective cross section of the nucleus to be transformed. By effective cross section is meant the circular area which a nucleus must have if *each* collision within this cross section would produce a transformation. It always depends on the velocity of the colliding particle. Since the probability of the occurrence of a reaction is contained in the concept of this effective cross section, the effective cross section is a theoretical concept which facilitates the computation of the yield, especially of thin bombarded foils or in referring to materials of different thicknesses. However, it has no direct relation to the actual nuclear radius or nuclear cross section as discussed on page 266. Nevertheless, for most nuclear reactions it lies between 10^{-23} to 10^{-27} cm^2, and has thus at least the order of magnitude of the actual cross section. Nuclear physicists have recently adopted their own unit for the effective cross section. They have given to the cross section 10^{-24} cm^2 the name "barn."

The effective cross section thus includes, first, the probability that a particle will penetrate into a nucleus; second, the probability that the transformation under consideration will actually occur. This second probability depends largely on the internal interaction of the nuclear particles. Classically, the probability that the colliding particle will penetrate the nucleus depends on the interaction time. For the case that the colliding particle is a neutron, the interaction time is inversely proportional to its velocity; it thus decreases with increasing velocity as $1/v$. For charged particles (α, p, d) the probability of penetrating into the nucleus increases exponentially at first with increasing impact energy because of the Coulomb repulsive force, i.e., the necessity of penetrating the potential wall of the nucleus. At higher energies it then decreases slowly because of the decreasing interaction time. A curve of the probability W of penetrating into the nucleus versus the particle energy is shown schematically in Fig. 147 for a neutron and for a proton.

This curve of the penetration probability would correspond directly to the effective cross section if, after the particle has penetrated into the nucleus, the probability of a transformation did not depend on the kinetic energy carried in by the colliding particle. However, such a dependency does, in general, exist. We have considered above (page 284) that nuclear transformations lead to the formation of intermediate nuclei, whose heating up by the kinetic and potential (binding) energies carried in by the colliding particle results in the emission of other nuclear particles, and that the life-time of the intermediate nucleus with respect to each decay determines its transformation probability. Only if the energy freed in the nucleus by the absorption of the colliding particle is exactly equal to an excitation step in the term diagram of the intermediate nucleus do we really have the intermediate nucleus, and then the transformation probability will be especially large; this is called energy resonance between the colliding particle and the intermediate nucleus. For these resonance energies the transformation probability and thus the effective cross sections suddenly become very large; for these energy values steep maxima are superimposed over the smooth curves shown in Fig. 147. These resonance energies are exactly known for a large number of nuclear reactions, especially with slow neutrons; and since they lead to direct conclusions about the energy levels of the nucleus, they aid greatly in supplementing our knowledge of nuclear structure. Also the principal features of our knowledge of the yield of artificially induced nuclear reactions thus can be explained rather satisfactorily in our pictorial theory of nuclear processes. Figure 146 shows the actual yield functions for some given nuclear reactions of the $_4Be^8$ nucleus with the pertinent transitions.

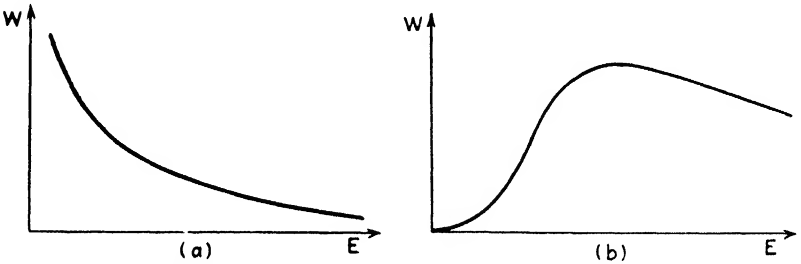


FIG. 147. Excitation functions of nuclear transformations (yield versus energy of the exciting particle), shown schematically, (a) for particles which are not repelled by the nucleus (neutrons) and (b) for positively charged particles which are repelled by the nucleus. Resonance cases not taken into consideration.

5-11. Nuclear Fission and the New Elements

In all the nuclear transformations, natural or artificial, which we have discussed so far, only individual particles (α, p, n) are ejected from the intermediate nuclei. In 1938, Hahn and Strassmann discovered a fundamentally different kind of nuclear transformation. They found that the radioactive products resulting from irradiating the rare uranium isotope $_{92}U^{235}$ with slow neutrons were not, as was first believed, so-called transuranium isotopes of atomic numbers larger than 92, but isotopes of well known elements of medium atomic numbers, such as barium. They concluded that these isotopes from the middle of the periodic table originated from a *splitting of the intermediate nucleus* $_{92}U^{236}$*, resulting from the absorption of a neutron by the* $_{92}U^{235}$ *nucleus, into two fractions with the average masses* $\frac{2}{5}$ *and* $\frac{3}{5}$ *of uranium* (see Fig. 148). This new process, the splitting of a heavy nucleus into two major fragments, now is called a *fission process,* the resulting nuclei, the *fission products.* Further investigation revealed that in each such fission process a large amount of energy, approximately 150 Mev, is released. As the intermediate uranium nucleus with 52 excess neutrons has a much greater neutron excess than have the stable nuclei of the fission products of medium mass numbers, the fission products must either transform neutrons into protons by multiple β-emission, thus as artificially radioactive nuclei transform themselves in steps into stable nuclei, or the excess neutrons must be ejected directly. Actually both processes take place. The fission products (see Fig. 149) actually have a considerable neutron excess which they reduce by several consecutive β-emissions; besides that, *in each fission process 3 to 4 neutrons are set free.* It seems that this neutron emission occurs at least partly during the fission process itself, partly by the highly excited fission fragments which cannot lose their neutron excess quickly enough by β-transformations.

FIG. 148. Cloud-chamber photograph of the fission of a uranium nucleus upon absorption of a neutron; the two fission products leave the horizontal uranium plate in opposite directions with a combined energy of approximately 150 Mev. (*After Corson and Thornton.*)

This emission of neutrons during and as a consequence of uranium fission is of utmost importance since these neutrons can induce fission in additional uranium nuclei and thus cause a chain reaction. Flügge seems to have been the first to point to this method of producing energy by a self-sustaining nuclear reaction, which was first achieved in 1942 by Fermi in Chicago. The uranium piles and the atom bomb, which are based on this principle, will be discussed on page 304.

The process of nuclear fission is not limited to the uranium nucleus $_{92}U^{235}$ or to neutron bombardment. It has been proved that a large number of nuclei of atomic number greater than 77 can be split by present experimental means. The energy necessary to induce fission varies considerably, corresponding to the different internal stabilities of the nuclei. The $_{92}U^{235}$ nucleus even splits spontaneously, although with only a small probability (long life-time), i.e., it requires no external energy at all, whereas its fission can be produced by thermal neutrons with very large cross section. Neighboring nuclei such as $_{92}U^{238}$ or thorium isotopes require faster, i.e., higher energy neutrons. Up to the present the lighter elements down to platinum can be split only by means of the 200-Mev neutrons or by 380-Mev α-particles produced by the large Berkeley cyclotron (Fig. 131). By these last remarks we have indicated that fission is also possible if the activation energy comes from other means than neutrons. In addition to neutrons, bombardment by other particles such as deuterons (which in hitting a nucleus mostly dissociate into a penetrating neutron and a deflected proton), α-particles, and finally high-energy photons (γ-rays) have proved effective. Because it is analogous to photoionization (page 107) and photodissociation (page 377) this last type of splitting is called *photofission*. According to work recently done in France, fission does not always produce two fragments (and several free neutrons) but *occasionally a third lighter fragment is ejected perpendicularly to the two main fragments which are hurled away diametrically from each other.* The mass of these lighter fragments averages about 5 with a maximum of 10; α-particles appear among them frequently. Rarely, with a probability of about 1 : 10,000, *quadruple fission* occurs.

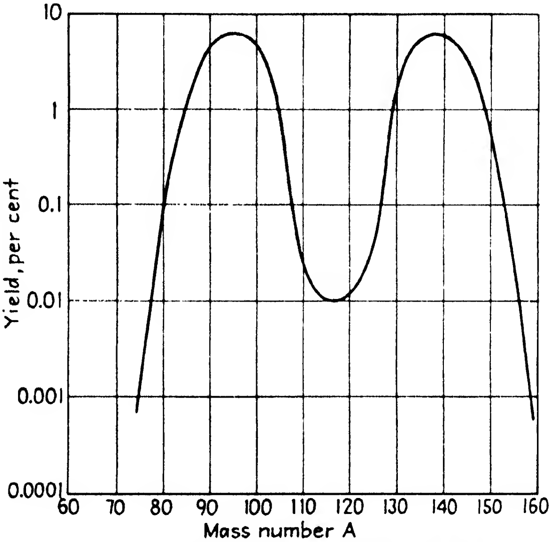


FIG. 149. The fission yield curve for neutron induced fission of $_{92}U^{235}$ nuclei (percentage of fission products versus mass number).

The masses of the normal fission products vary within very wide limits; indeed almost all the elements of the periodic table are found among them, although in very different abundance. From the abundance curve, Fig. 149, it can be seen that there are two steep maxima for the mass numbers near 95 and 140, which means that a fission in that mass ratio of 40 to 60 per cent is much more probable than a splitting into two fragments of approximately the same mass. However, this asymmetry of fission seems to be characteristic only for excitation by slow neutrons and disappears for excitation by particles of higher energy. For fission of uranium by 380-Mev α-particles from the Berkeley cyclotron a yield curve with only one maximum at 115 has been found, and a similar curve with a maximum at a mass number of about 100 was obtained for the fission of Bi by 200-Mev neutrons.

As a result of intensive research with nuclear fission, excitation functions (see page 299), i.e., the relation between the fission probability and the energy of the particles which induce the fission, have also recently been measured directly. Deuterons and α-particles were used in these first experiments, because it is more difficult to determine the energy of neutrons than that of particles which can be accelerated electrostatically. The quantitative evaluation of the experiments shows the effective cross section to be up to 3×10^{-24} cm^2 (3 barn) which agrees with the actual cross section of the nucleus.

Theoretically the splitting of very heavy nuclei by adding energy is not too surprising. That the nuclei of the heavier elements of the periodic table are somewhat unstable even without any activation energy is indicated by the phenomenon of natural radioactivity. If we consider, furthermore, that according to page 267 the heavy nuclei show a definite deviation from the normal spherical shape, it does not seem too hard to understand that a transfer of energy can cause internal vibrations and thus finally a splitting of these nuclei. Computations based on the liquid-drop model, used by von Weizsäcker to explain nuclear binding, confirm this conclusion. According to the liquid-drop analogy the nucleus is held together by forces which act like a surface tension, whereas the electrostatic repulsion between the protons tries to break up the nucleus. Computations by Bohr and Wheeler, based on Eq. (5-36), have led to the result that for nuclei with atomic numbers above $(47A)^{\frac{1}{2}}$, where A is the mass number of the nucleus, the internal electrostatic pressure between the protons should exceed the surface tension which is caused by the specific nuclear forces. Actually nuclear fission has been observed for nuclei with even somewhat smaller atomic numbers, if they are excited by some activation energy. This is caused by internal mechanical vibrations within the nucleus which become unstable and lead to its fission. Dänzer has been able to demonstrate such vibrations very nicely in model experiments with soap bubbles whose surface tension was nearly compensated by an electric surface charge. The splitting of the bubble could be observed here, too.

Computations on the basis of Eq. (5-36) lead to the result that the fission of one of the heaviest known nuclei into two equal halfs due to internal electrostatic repulsion would cause the release of approximately 300 Mev energy. However, half of this amount is used up for the resulting increase of the surface, so that a net energy of 150 Mev should be released per fission process according to the liquid-drop model. This is in excellent agreement with observation. Thus nuclear fission can be easily understood on the basis of this simple model. The fact that it is possible to split lighter nuclei also, though with a large activation energy, fits well into the picture. Whereas neutrons with 6 to 8 Mev binding energy are sufficient to split U^{235}, much greater activation energies, which can be furnished only by accelerated exciting particles, are required to split the more stable nuclei.

Since, according to recent research, ${}_{92}U^{235}$, with a very small probability and the correspondingly long half-life of about 10^{14} years, can split spontaneously, *the idea seems reasonable that the process of fission is responsible for the fact that the periodic table breaks off with uranium.* Thus a much discussed fundamental problem of atomic physics seems to have found its solution.

The breaking off of the periodic table as a result of increasing instability due to α-decay, β-decay, and fission is not an abrupt one. Therefore, it is not surprising that the irradiation of uranium with neutrons yields isotopes of at least four "transuranium" elements with atomic numbers 93 to 96. These new elements were named *neptunium*, ${}_{93}Np$; *plutonium*, ${}_{94}Pu$; *americium*, ${}_{95}Am$; and *curium*, ${}_{96}Cm$. Table 11 shows the presently known isotopes of the transuranic elements with their decay processes and half-life times.

The most important of these new nuclei is the *plutonium* isotope ${}_{94}Pu^{239}$ because it can be split as easily as ${}_{92}U^{235}$ when bombarded with slow neutrons. We shall return to the technical importance of this process later. As the result of adding neutrons to the most abundant uranium isotope the following reaction occurs:

$$ {}_{92}U^{238} + n \rightarrow {}^{-}_{92}U^{239} \rightarrow {}^{-}_{93}Np^{239} + e^{-} \rightarrow {}_{94}Pu^{239} + e^{-} \qquad (5\text{-}27) $$

The ${}_{92}U^{238}$ thus is transformed to a β-active uranium isotope 239 which, by double electron emission (naturally always coupled with neutrino emission), transforms into neptunium and finally plutonium.

The third easily fissionable nucleus, besides U^{235} and Pu^{239}, is the formerly unknown uranium isotope U^{233}. In a reaction similar to that which produces plutonium, it is produced as the final result of the bombardment of thorium by neutrons,

$$ {}_{90}Th^{232} + n \rightarrow {}^{-}_{90}Th^{233} \rightarrow {}^{-}_{91}Pa^{233} + e^{-} \rightarrow {}_{92}U^{233} + e^{-} \qquad (5\text{-}28) $$

Table 11. Isotopes of the Transuranium Elements

Element	Mass number	Type of decay	Half-life
Neptunium, Np.....	231	α, K	53^{m}
	234	K, γ	4.4^{d}
	235	K, α	435^{d}
	236	β^{-}, γ	22^{h}
	237	α	2.20×10^{6y}
	238	β^{-}, γ	2.0^{d}
	239	β^{-}, γ	2.33^{d}
Plutonium, Pu......	232	α	22^{m}
	234	α, K	8^{h}
	236	α	2.7^{y}
	237	K	40^{d}
	238	α	90^{y}
	239	α, γ	2.4×10^{4y}
	240	α	$6{,}000^{y}$
	241	β^{-}, α	10^{y}
Americium, Am.....	238	K	1.5^{h}
	239	K	12^{h}
	240	K, γ	50^{h}
	241	α, γ	500^{y}
	242	α, β^{-}	400^{y}
Curium, Cm........	238	α	2.5^{h}
	240	α	26.8^{d}
	241	K	55^{d}
	242	α	150^{d}

5-12. The Liberation of Useful Nuclear Energy

We mentioned above that the discovery of neutron emission upon nuclear fission has made it possible to liberate nuclear energy, either in the catastrophic form of the exploding nuclear bomb, or controlled in the nuclear piles which can continuously produce energy. Because of this fact, nuclear and atomic physics suddenly have become the center of general interest as well as of political discussion. Unfortunately the term "atomic energy," instead of the correct term "nuclear energy," has come into common use and has caused an apparently irreparable confusion among laymen as to the difference between atomic and nuclear processes. In our presentation we shall discuss only the most essential features of the so-called atomic bomb and those of the controlled release of nuclear energy.

The mechanism of the nuclear bomb is easily understandable. If a mass of pure U^{235}, U^{233}, or Pu^{239}, i.e., of those isotopes which are extremely fissionable when bombarded by neutrons, has a diameter larger than the mean free path of a neutron released in the metal, *one* slow neutron or *one* spontaneous nuclear fission will produce an explosion, because with each fission several neutrons are released, each of which in turn causes further splitting, thus producing an avalanche (chain reaction) of splitting nuclei. For each ${}_{92}U^{235}$ nucleus that is split, 150 Mev are released. Thus, if every nucleus in one pound of U^{235} is split, the energy released is 3×10^{20} ergs = 8 million kwh = 7×10^9 kcal.

One of the outstanding accomplishments in the development of the atomic bomb was the theoretical computation of the material dimensions for a bomb. The computations of the mean free path of neutrons in the bomb material were based on experimental investigations with microscopic quantities of U^{235} or Pu^{239}. Naturally this information is still highly secret.

One of the technical difficulties in producing the uranium bomb (we shall come to the plutonium bomb later) was the purification of the fissionable isotope U^{235}, i.e., its separation from the nonfissionable U^{238}. According to Table 3, page 59, natural uranium consists of only 0.72 per cent of the fissionable isotope U^{235}. The difficulty was overcome, but the technical expenditure for the large-scale separation was tremendous compared even with that required for the large-scale production of plutonium. It seems doubtful, therefore, whether pure U^{235} will be produced in the future.

The second difficulty in constructing any nuclear bomb is prevention of its spontaneous explosion. This spontaneous explosion is initiated immediately by *one* spontaneous fission process or a stray neutron if an amount of fissionable material larger than the "critical mass" is assembled at one spot. The only way to prevent spontaneous explosion of the bomb is to keep it in two pieces, each less than the critical size, and to join these at the moment the initiation of the explosion is desired. This joining of the two parts of the bomb must take place so rapidly and completely that the greatest possible number of nuclei are split before the bomb explodes mechanically as a result of the intense internal production of energy, thus stopping the reaction. Therefore, the amount of energy which we computed above to be liberated by fission is only a theoretical upper limit and does not allow conclusions as to the amount of energy actually released in the explosion of a nuclear bomb.

In contrast to the bomb, the uranium pile continuously produces an adjustable amount of energy; thus the number of nuclei split per unit of time must be constant. Therefore, a mixture of the fissionable isotope 235 with the "inactive" isotope 238 is used as raw material, the percentage of 235 varying with the amount of energy to be produced. The splitting is regulated (by means yet to be discussed) so that per unit time a fixed number of neutrons induce new fission processes while part of the excess neutrons are absorbed by the U^{238} nuclei. As mentioned above, these more stable nuclei are not split by slow neutrons, but the process leads to the formation of plutonium Pu^{239} which is produced according to the reaction (5-27). A certain number of neutrons are scattered and absorbed by other inactive nuclei. Since slow neutrons are required for splitting U^{235} just as for Pu^{239}, the fast neutrons produced during fission must be retarded, i.e., their energy must be dissipated by scattering. For this, a scattering material of small atomic weight is required, since according to page 291 the amount of energy transferred in a neutron collision is the larger, the smaller the atomic weight of the scattering atom. The scattering material, furthermore, must have the smallest possible absorption coefficient for neutrons so that as few valuable neutrons as possible get lost by absorption. Deuterium in the form of heavy water, D_2O, as well as purest carbon in the form of graphite can be used as retarding substances (so-called moderators). Graphite actually is used more frequently because it is less expensive. To summarize: In the pile, Pu^{239} is produced from U^{238}, while simultaneously fissionable material, U^{235} *and* Pu^{239}, is lost due to fission processes. However, since the gain of Pu^{239} may be greater than its loss due to fission, the concentration of fissionable material in the uranium can increase with time. Consequently, the energy production in this arrangement would increase steadily and finally lead to its destruction if the number of fission processes could not be regulated. This can be done in a theoretically neat and practically simple manner. Since cadmium, according to page 292, has a large cross section with respect to the absorption of slow neutrons, it can be used to capture neutrons and thus reduce the number of neutrons which are available for splitting nuclei.

A uranium pile (Fig. 150) thus consists of a large number of uranium rods which are enclosed by metal casings to prevent surface corrosion and are imbedded in neutron-retarding graphite. Cadmium rods are provided which can be inserted more or less deeply into the pile to regulate the production of energy and, if necessary, can be used to shut off the machine. A cooling system removes the heat generated from the pile. The cooling system operates with a cooling fluid (possibly a liquid metal) or a gas (helium). In practice the absorbed heat is either converted into superheated steam which can be used to operate a steam turbine, or operates a gas turbine directly. The dimensions of the various piles operating at the present time in the United States, Canada, England, and France are not known. However, it seems that these piles are huge and heavy structures, particularly so because thick concrete walls are necessary for protection against radiation. Considerable experience with nuclear energy and entirely new principles of construction will be necessary, therefore, before nuclear-energy machines will be used to operate ships, airplanes, or other vehicles.

If it is desired to use a pile not only for production of energy but also for the production of plutonium (as bomb material or as fuel for high-power piles), the uranium rods are removed after they have been sufficiently enriched with Pu^{239} and the plutonium is separated by chemical methods. Because of the intense radioactive radiation of the fission products all manipulations with the rods, their removal from the pile, and their chemical processing must be carried out by remote control. Furthermore, equipment used in these manipulations must be guarded so that human beings do not come in contact with it. Consequently, a plutonium plant presents quite a technical problem. The same is true for the production of fissionable U^{233} from thorium, using the process (5-28) discussed on page 303.

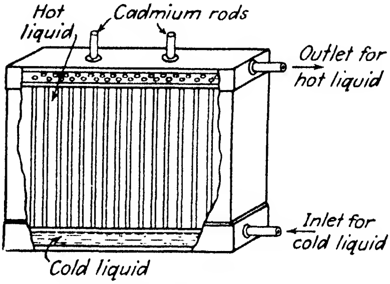


FIG. 150. Schematic diagram of a uranium pile: uranium rods in a moderator, cooled by a liquid.

It seems obvious, from all this, that any large-scale industrial utilization of fission energy depends on the economical solution of many problems of safely processing radioactive materials. There is the development of suitable materials, especially for building piles and cooling systems. These materials must combine good mechanical properties at high temperatures with other specific properties (behavior with respect to neutrons, corrosion, thermal conductivity, etc). Another problem which must be solved is the economical "breeding" of new fuel (Pu^{239} or U^{233}) and its separation from the nonfissionable material, as well as the safe disposal of the radioactive waste products. Opinions about the time required to solve all these problems differ widely.

The nuclear piles are, besides their possible industrial use, of great scientific interest for two reasons. First, intense neutron beams can leave the pile through "windows" in the concrete shields and can thus be used for experimental investigations. The intensity of these neutrons exceeds by many orders of magnitude the intensity of any other known neutron source. Of still greater significance, however, is the second application. As a result of the fission processes, radioactive fission products embracing the major part of the periodic table accumulate in the uranium rods (see Fig. 149). Among them are a large number of isotopes which are important for use in physics, chemistry, biology, and medicine. Some of these isotopes can be separated chemically and then are sent all over the world to interested scientific institutions. Such radioactive isotopes can also be produced by intensive neutron irradiation of samples inserted into the pile [(n,γ) processes, page 292]. In addition to the transuranic elements Np, Pu, Am, and Cm the "missing" elements in the periodic table, 43, 61, 85, and 87 are among the isotopes which have been isolated from the pile. These have been separated in sufficient quantities so that their chemical properties could be investigated by the methods of microchemistry. They are naturally radioactive and for this reason no longer occur in a natural state on the earth. These new elements have been named technicium, $_{43}$Tc, promethium, $_{61}$Pm, astatine, $_{85}$At, and francium, $_{87}$Fr, by their discoverers.

In conclusion, we call attention to the fact that uranium fission, in spite of the large energies released by each process, is not the most exothermal nuclear process known, if we refer to the production of energy per unit mass. According to the process (5-26) the energy 17.2 Mev is released when lithium $_3\mathrm{Li}^7$ decays into two α-particles upon bombardment by protons. This is only 11.5 per cent of the energy of uranium fission; however, since the lithium nucleus has only 3 per cent of the mass of the uranium nucleus, the energy production per mass unit of transformed material according to process (5-26) is four times as great as that of the fission of uranium 235. However, the technical utilization of this process does not yet seem to be possible, because no chain reaction involving this nuclear process is known.

5-13. Further Applications of Nuclear Physics, Especially of the Radioactive Isotopes

Now that we have discussed the technical utilization of nuclear fission we shall make a brief survey of further applications of nuclear physics, especially of radioactive isotopes.

In addition to the production of energy by nuclear fission there is another possibility of obtaining nuclear energy by a reverse process, the synthesis of He^4 nuclei from protons according to (5-31) with the liberation of about 26 Mev. We shall see in the next section that solar energy, which makes life possible on this earth, originates from these thermal nuclear-synthesis processes which occur in the interior of the sun as well as in most fixed stars.

In general, these processes which occur in the interior of the sun cannot be used with profit in the laboratory because by far the largest part of the colliding particles lose their energy in ionizing collisions with shell electrons and only seldom, according to page 297, does such a particle collide with a nucleus and then cause a transformation. For this reason, the energy used to accelerate the colliding particle exceeds the possible energy gain from artificial exothermal nuclear reactions by many orders of magnitude. *Only because in nuclear fission several neutrons are released in each fission process which can excite other nuclei and thus cause a chain reaction, is nuclear fission more profitable for producing energy.*

We leave the question of the liberation of nuclear energy, which has disturbed the whole world so much (and which unfortunately seems to be the only interesting feature of atomic physics so far as the general public is concerned), and consider the scientific, technical, and medical applications (in the broader sense) of nuclear physics. The fact that the gaps in the periodic table were closed by the discovery of the *previously unknown elements* ${}_{43}$Tc, ${}_{61}$Pm, ${}_{85}$At, and ${}_{87}$Fr, and the transuranium elements ${}_{93}$Np, ${}_{94}$Pu, ${}_{95}$Am, and ${}_{96}$Cm, has been mentioned already as an application of nuclear physics. The nuclear theory, to be discussed on page 321, tells us that stable isotopes of all these elements do not exist; thus they are all radioactive. Nevertheless, all these elements, in the meantime, have been isolated by Seaborg and others to such an extent that their chemical properties (often with samples of only a few micrograms) have been tested very precisely. This is a truly remarkable achievement of microchemical experimental technique. It is especially interesting, as mentioned on page 176 and indicated in Fig. 99, that the elements 90 to 96 show a very similar behavior. There is no doubt any more that these elements constitute a closed group, similar to the rare earths. This is caused by the fact that these elements have the same external electron arrangement and differ only in the number of their inner $5f$ electrons (see page 179).

In addition to the stable isotopes (especially the heavy hydrogen isotope), artificially radioactive atoms have become increasingly important as indicators or tracers, for they make it possible to identify and follow a particular atom among similar atoms through an entire reaction by means of its radiation. The incorporation of foreign atoms into a crystal lattice (see page 456), for example, the application of radioactive foreign atoms, made it possible to show that the foreign atoms prefer to be at the boundary between microcrystals and not in the lattice itself. The self-diffusion of a substance in a crystal built up of identical atoms can also be measured by means of this tracer method, in which the progress of the diffusing atoms can be followed by checking their β-radiation. In a similar way, the speed of reaction and the reaction mechanism can be successfully studied by the radioactive-tracer method. To cite only one interesting example, this tracer method has helped in clearing up a question which is of high importance for botany: the mechanism of the assimilation of carbon dioxide by green plants, which can be studied in its various stages by this method. With regard to the importance of carbon in organic chemistry it is of interest that there are three carbon isotopes which are well suited as radioactive tracers — the two positron radiators ${}^{+}_{6}\mathrm{C}^{10}$ and ${}^{+}_{6}\mathrm{C}^{11}$ with half-lives of 19 sec and 21 min, respectively, and the electron radiator ${}^{-}_{6}\mathrm{C}^{14}$ with a half-life of 10^3 years. Because of its convenient half-life, ${}^{+}_{6}\mathrm{C}^{11}$ is very important. However, for experiments lasting over a long period of time, the long-lived ${}^{-}_{6}\mathrm{C}^{14}$ is more suitable if it is possible to work with a sufficient concentration of active C^{14} so that measurements are not too difficult.

The indicator method is also used with great success in biology and medicine. The speed of migration of certain substances in the body, when absorbed with food, injected, or introduced by other means, has been studied in this way, and it is surprising how quickly an exchange of substances occurs inside the body. Also, the normal metabolism and the functioning of the different individual organs in normal as well as pathological cases can be studied by the tracer method. Thus, in the latter case, nuclear physics serves as a direct diagnostic aid to the physician. Of special importance for all these medical experiments is the fact that there exist some radioactive atoms which are not foreign to the human body and thus do not produce any disturbing secondary effects. Among these are the electron radiators, ${}^{-}_{11}\mathrm{Na}^{24}$ (radiosodium) with a half-life of 15 hrs and ${}^{-}_{15}\mathrm{P}^{32}$ (radiophosphorus) with a half-life of 14.1 days, which have suitable half-lives for experiments. Since different active atoms have a different preference for depositing in bones, tissues, and organs, there is at least a possibility that they may be used to exert specific therapeutical influences. It is certain that this field will be extensively investigated in the future and will find wide application. As an interesting example of this type, mention may be made of the fact that radioactive phosphorus by its β-emission has a considerable influence on the concentration of the different blood constituents and has been used with some success in treating a special blood disease, leukemia.

The application of nuclear physics in technology and engineering is still in its infancy, but it may be mentioned that radioactive compounds emitting hard γ-rays, and recently also betatron-produced very hard X-rays, are used to search for casting flaws and other types of defects in thick metal pieces which normal X-rays cannot penetrate.

The extent of all these and many other applications will increase rapidly in the near future.

5-14. Thermal Nuclear Reactions in the Interior of Stars at Extremely High Temperatures. The Origin of the Elements

The nuclear reactions which we have discussed so far are induced, if we disregard the spontaneous decay of naturally radioactive elements, by bombarding the nucleus with naturally or artificially accelerated particles. How do the normal chemical reactions compare with this? These can also be initiated by individual accelerated particles (electrons or ions). However, they usually occur "in thermal equilibrium." For instance, by heating oxyhydrogen gas the average velocity of its molecules increases until, according to Maxwell's velocity distribution, some molecules obtain sufficient energy to dissociate and thus initiate a reaction (explosion). Our comparison leads to the question whether there are also thermal nuclear reactions which are initiated by fast nuclear particles that obtained their kinetic energy as a result of the high temperature of the gas (plasma) of which they are a part.

Such thermal nuclear reactions do exist, but they are not easily obtainable because, in comparison to chemical reactions, they require such large "activation energies" (in order to enable a particle to penetrate into a nucleus and thus initiate a reaction) that the temperature must be many orders of magnitude higher than any we can produce on earth. This temperature is of the order of magnitude of 10^7 degrees.

According to the fairly reliable computations of the astrophysicists, such temperatures exist in the interior of the fixed stars, e.g., the interior of the sun has a temperature of about 20 million degrees. C. F. von Weizsäcker first pointed out in 1936 that exothermal nuclear reactions must be possible at these temperatures and that these thermal nuclear reactions, which release huge amounts of energy, were responsible for the origin of the permanent radiation of the sun. A long-standing problem thus found its solution, because astronomers had tried in vain for decades to find the origin of the huge amount of energy which our sun has emitted constantly for at least many millions of years. Von Weizsäcker's hypothesis has been confirmed, and from further work, especially of Gamow and Bethe, we have a great deal of detailed knowledge of the particular nuclear reactions which go on in the interior of the sun.

First of all we can compute the average kinetic energy $(m/2)v^2$ of a particle for the central temperature of the sun, 2×10^7 degrees, from the equation

$$\frac{mv^2}{2} = \frac{3}{2}kT \tag{5-29}$$

It turns out to be about 2,600 ev. This small value for the average particle energy, compared to the millions of electron volts which we use in nuclear physics laboratories, is compensated, first, by the great number of collisions which make even extremely improbable reactions sufficiently frequent, and, second, by the fact that, according to Maxwell's distribution of velocities, a small number of particles have average energies far in excess of the average energy $\frac{3}{2}kT$. These relatively few particles, therefore, are able to initiate the necessary number of nuclear reactions. Since the number of particles having sufficiently high velocities, and the probability of a particle's penetrating into the nucleus increase with the temperature, we have a very rapid increase in the number of reactions per second and per cubic centimeter with the temperature, or in the technical term of chemistry, we have a high temperature coefficient of thermal nuclear reactions.

From the experimentally known yield of most nuclear reactions and the half-life of the radioactive isotopes, the thermal reactions possible at the sun's central temperature of 2×10^7 degrees can be computed fairly accurately. This work of Bethe and Gamow led to highly interesting results. First of all it shows that all light nuclei up to and including ${}_5B^{11}$ react very quickly (compared to cosmic periods of time) with the very abundant protons in the sun and fixed stars and, in different ways, they always build up the extremely stable helium nuclei ${}_2He^4$ (α-particles). This result seems to explain why the light nuclei (and elements) ${}_1H^2$ up to and including ${}_5B^{11}$ are so rare in the sun and on the earth. These computations also show that *the heavier nuclei, from* ${}_8O^{16}$ *upward, remain practically unchanged for periods up to* 10^{12} *years and temperatures up to* 2×10^7 *degrees.* This is understandable since neutrons apparently are not produced in the sun, and the protons with their positive charge cannot overcome the repulsion of the high positive charges of the heavy nuclei and penetrate into them. *Thus a variation in the distribution of the heavy elements from oxygen on up is not very probable in the sun and the fixed stars.*

Of special importance, however, are the reactions of the nuclei ${}_6C^{12}$ to ${}_7N^{15}$ with protons and the subsequent secondary reactions. They can be written, according to Bethe, as follows:

$$\left.\begin{array}{lll}
{}_6C^{12} + p & \longrightarrow & {}^{+}_{7}N^{13} \\
{}^{+}_{7}N^{13} & \longrightarrow & {}_6C^{13} + e^{+} \\
{}_6C^{13} + p & \longrightarrow & {}_7N^{14} \\
{}_7N^{14} + p & \longrightarrow & {}^{+}_{8}O^{15} \\
{}^{+}_{8}O^{15} & \longrightarrow & {}_7N^{15} + e^{+} \\
{}_7N^{15} + p & \longrightarrow & {}_6C^{12} + {}_2He^4
\end{array}\right\} \tag{5-30}$$

The normal carbon nucleus ${}_6C^{12}$ reacts with a proton to produce the positron-active nucleus ${}^{+}_{7}N^{13}$ which goes over into ${}_6C^{13}$ by emitting a positron. ${}_6C^{13}$ reacts again with a proton to produce the stable ${}_7N^{14}$, and this in turn reacts with another proton to produce the radioactive ${}^{+}_{8}O^{15}$, which then reverts to ${}_7N^{15}$ by emitting a positron. The nucleus ${}_7N^{15}$ finally reacts with a proton and decays thereby [(p,α) reaction] to the original stable nucleus ${}_6C^{12}$ and an α-particle, ${}_2He^4$.

The surprising and important result of this reaction cycle is, thus, that the original nucleus ${}_6C^{12}$ is not used up, but is reproduced at the end of the cycle. Furthermore, in the cycle four protons combine to form a He^4 nucleus, releasing two positrons, so that the balance of the reaction cycle can be written

$$4\,{}_1H^1 \rightarrow {}_2He^4 + 2e^+ \tag{5-31}$$

This reaction is, as we know from page 271, strongly exothermic, i.e., energy is released, because the mass of four protons equal to 4 × 1.0076 is 0.0275 mass unit greater than the resulting ${}_2H^4$ with 4.0029 mass units. Therefore in this reaction cycle the large amount of energy

$$\begin{aligned} 25.6 \text{ Mev per He nucleus} &= 1.5 \times 10^8 \text{ kcal/gm helium} \\ &= 6.0 \times 10^8 \text{ kcal/mole} \end{aligned} \tag{5-32}$$

is released. In the language of the chemist, (5-31) is a catalytic reaction with ${}_6C^{12}$ as catalyst, since the carbon nuclei come out of the reaction unchanged, and carbon thus is not expended.

Exact computations show that *by this thermonuclear reaction cycle, the catalytic transformation of four protons into a helium nucleus, sufficient energy is supplied to make up for the total permanent radiation of the sun.* The sun's supply of hydrogen, furthermore, is sufficient to last for approximately 10^{11} years.

It may be assumed that in the other normal fixed stars, i.e., those in the principal series in the Hertzsprung-Russell diagram, the energy is also produced by this reaction cycle. On the other hand, it is not yet clear how the energy is produced in the abnormal cases, the white dwarfs and the giants. At the much higher temperatures and pressures which prevail in the interior of the white dwarfs and (according to von Weizsäcker) also in the interior of the nuclei of the giants, it is possible that, to a small extent, neutrons are also produced by collisions of protons with ionized shell electrons according to the equation

$${}_1H^1 + e^- \rightarrow {}_0n^1 \tag{5-33}$$

According to a new (and as yet unconfirmed) idea of Jordan, all giant stars should have a compact neutron core originating from the process (5-33), which, in connection with the radioactive decay of neutrons mentioned on page 289, tends to cause pulsations and might play a role in the mechanism of the pulsating stars. In general, however, by far the majority of these neutrons react first with hydrogen (protons) to form heavy hydrogen ${}_1H^2$ and eventually ${}_2He^4$, thus producing nothing more than does the normal reaction cycle (5-30). Because of the large excess of hydrogen in the stars compared to the heavy elements, only a few neutrons are absorbed by heavy nuclei before they are captured by protons. Thus they have little opportunity for building up higher heavy nuclei.

At extremely high temperatures (of the order of 10^{10} degrees), which may be temporarily reached in certain stages of development of stars, another interesting nuclear process, according to a hypothesis of Gamow, should play an important role: a stable nucleus such as ${}_8O^{17}$ absorbs electrons in a manner similar to the *K* effect (page 282) (however, this time the electrons are taken not from the shells but from the electron gas in the vicinity of the nuclei) and simultaneously emits an antineutrino ν_a.

$${}_8O^{17} + e^- \rightarrow {}_7N^{17} + \nu_a \qquad (5\text{-}34)$$

However, since the resulting nitrogen is electron-active, it immediately reemits an electron and a neutrino ν,

$${}_7N^{17} \rightarrow {}_8O^{17} + e^- + \nu \qquad (5\text{-}35)$$

so that the ${}_8O^{17}$ is reproduced and the process can begin anew. However, the two emitted neutrinos carry energy away with them, and because the interaction between neutrinos and matter is negligible, the neutrinos are practically unabsorbed and therefore can leave even the most massive stars. This strange process thus would result in an appreciable energy loss (cooling process), which had not hitherto been taken into account. According to Gamow's computations, the huge amount of energy of 10^{17} ergs could be carried away per gram of material per second.

It is not impossible that this process plays a role in the case of a star which has consumed all its hydrogen in the normal cycle (5-30) and then contracts because of gravitational forces, thus causing a tremendous increase in its internal pressure and temperature. The identification of the result of this neutrino process with certain observed unstable star types (supernovae), however, is still in the realm of speculation, and there are a number of other explanations for these remarkable stars.

According to the above mentioned calculations of nuclear reactions, the building up of heavy nuclei beyond neon is highly improbable during the "normal" life of a star, so that the problem of the origin of elements in their present distribution on the sun and earth (which seems to correspond also to that in the fixed stars) is still open. To our knowledge the highest temperatures which exist, at least in normal stars, do not exceed 100 million degrees, so that an appreciable shift in the present distribution of the abundance of heavy elements above oxygen is improbable. On the other hand, *there is no doubt that the elements in their present abundance must have been built up of protons and neutrons. This process of building up, which must have occurred at much higher temperatures than* 10^8 *degrees, must then have been broken off by a sudden decrease in temperature.*

It was first suggested by von Weizsäcker that one might think of some kind of primary explosion of our whole world, i.e., the totality of our fixed stars and nebulae. This process would have given "birth" to our present world and could easily have produced the high temperatures required to build up the heavy nuclei. Bethe and Gamow have pursued this idea recently and have assumed that the starting point of the entire universe was a highly compressed neutron gas at an extremely high temperature. By the radioactive decay of neutrons, discussed on page 289, protons and electrons were produced, then deuterons by combination of neutrons and protons. By further combination with neutrons, followed by stabilizing β-decay, all the other heavy elements would have been produced, until the gradual cooling, caused by expansion, brought an end to this process of building nuclei, allegedly in a few minutes. Certain consequences of this hypothesis seem to be in remarkable agreement with experience. For example, the relative abundance of the different isotopes (nuclei) is determined not so much by their own stability as by their ability to absorb neutrons. It still remains to be seen what future research will produce from these interesting, but highly speculative, hypotheses.

The determination of the age of radioactive elements from the concentration of their decay products in minerals on our earth as well as from outer space (meteors) indicates that the time of such an original explosion should have been several 10^9 years ago. In remarkable agreement with this, observations of astrophysics (e.g., the red shift in the spectra of the spiral nebulae and the expansion of the universe which produced it) indicate that *the state of our universe approximately* 3 *to* 7 × 10^9 *years ago was essentially different from what it is today.* Most fascinating problems here still wait for their solution. It is interesting to see from this example *how research on the smallest material particle, the nucleus, can give us a hint as to the fate of the universe as a whole.*

Gamow recently has extended these ideas somewhat more quantitatively. He assumed that for the indicated building-up process the temperature was so high (10^9 to 10^{10} degrees) that the average thermal energy was of the same order of magnitude as the average binding energy of the nucleon in the nucleus (8 Mev). At such high temperatures, however, the radiation density is so high, that the radiation pressure far exceeded the gravitational attractive forces of the existing matter and thus caused a rapid expansion of the universe. When then, as a consequence of the expansion, the radiation pressure had decreased until it was comparable to the gravitational forces, local condensations could occur in the uniformly expanding universe. The masses and diameters of these condensations have been computed by Gamow. The values obtained are in surprising agreement, at least in order of magnitude, with those known for the galaxies.

5-15. Structure and Systematics of Atomic Nuclei

We have already mentioned that a nucleus can be compared, in good approximation, to a drop of liquid. In support of this concept we mention, first, that the density of all nuclei is approximately constant, and, second, that according to Fig. 152 the binding energy of the nuclei per nucleon is almost constant and equal to 8 Mev. It decreases slowly to 7 Mev for the heaviest elements. From these two results we must conclude that, in contrast to electrostatic forces, the binding nuclear force has such a short range that it extends no farther than to the nearest neighbor. This conclusion can be experimentally verified. If neutrons are shot against protons, or protons and neutrons against similar particles, then, by measuring the angular distribution of the scattering, a conclusion can be drawn about the range of the nuclear forces. If we do not consider the force itself, but its potential, then we can, as we have done for atomic binding (see Fig. 119), plot this potential as a function of distance between the two particles under consideration. Thus we get the $U(r)$ curves, Fig. 151, for the proton-neutron and proton-proton potential. They differ in that the proton-neutron potential has no positive (repulsive) contribution. The nuclear forces, whose nature we shall discuss on page 335, have, according to these measurements, the small range which we expected. They are practically nonexistent beyond 5×10^{-13} cm.

One more important point can be gathered from Fig. 151. At first one might think that the large energy corresponding to the minimum of the $U(r)$ curve would be released as binding energy when the two particles combine and thus, conversely, must be expended to separate them. Actually the binding energy amounts to only about 10 per cent of the energy of the potential minimum, while the particle retains the remaining 90 per cent as kinetic energy in the bound state. This follows from the uncertainty principle (page 198). Since the particle is limited to the small volume of the nucleus, it has a correspondingly large uncertainty of momentum, i.e., velocity, from which the maximum kinetic energy of the particle, $(m/2)v^2$, can be computed.

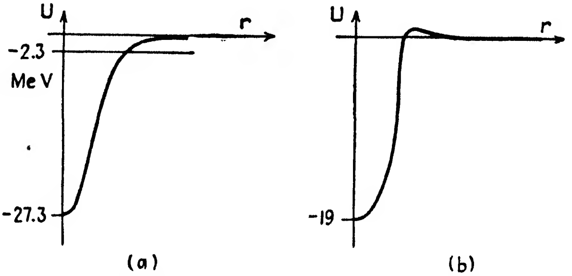


FIG. 151. Interaction potential $U(r)$ between proton and neutron; (a) no electrostatic repulsion; and between two protons (b) electrostatic repulsion superimposed on exchange attraction. Ordinate in million electron volts. (*After Heisenberg.*)

Thus the basis of our liquid-drop model seems to be a fairly sound one, and we can draw a number of detailed conclusions and thus determine the exact relation between the binding energy per particle and the atomic weight (Fig. 152). According to von Weizsäcker the total

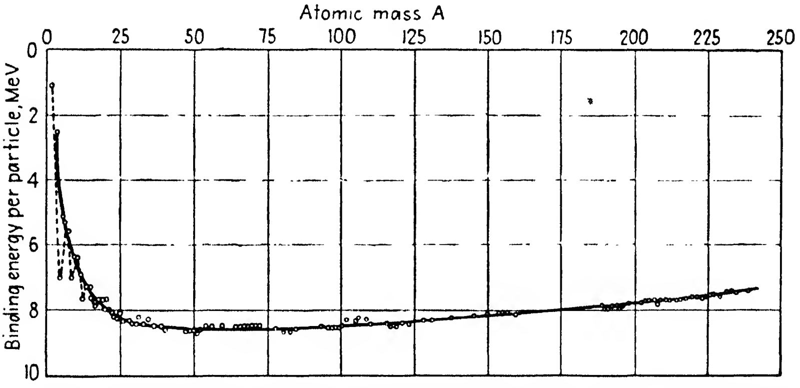


FIG. 152. Binding energy per nucleon (in million electron volts) plotted against the mass number of the whole nucleus.

energy of a nucleus, and thus also the binding energy per particle, is composed of four parts:

1. A constant contribution due to the nuclear forces between paired protons and neutrons

2. A slightly repulsive contribution of the surplus neutrons in the nucleus

3. A kind of surface tension due to the unilateral binding of the surface nucleons
4. The mutual repulsion of the protons

If N is the number of neutrons and Z the number of protons, we have for the binding energy per particle

$$\frac{E}{N+Z} = -A + B\frac{(N-Z)^2}{(N+Z)^2} + \frac{C}{(N+Z)^{\frac{1}{3}}} + \frac{3}{5}\frac{Z^2e^2}{(N+Z)^{\frac{4}{3}}r_0} \qquad (5\text{-}36)$$

For light nuclei, the fourth term with Z^2 can be neglected. We then have greatest stability if the second term vanishes, i.e., for $N = Z$. Actually the most stable light nuclei have an equal number of protons and neutrons. The small average binding energy per nucleon of the light nuclei (see Fig. 152) evidently is due to the fact that a relatively large fraction of nucleons are surface nucleons and thus not so strongly bound as inner nucleons. This is expressed in the third term of (5-36). In the heavy nuclei the term with Z^2 depending on the Coulomb repulsion becomes more effective and is responsible for the slow decrease in binding energy per particle. It is the cause of the α-emission of the heavy naturally radioactive nuclei. Because of the value of the fourth term for heavy nuclei, it is energetically favorable for a nucleus to have an excess of neutrons (in agreement with experience). The relative increase of the second term of (5-36) is then overcompensated by the decrease of the fourth term. Thus this formula, derived from the pictorial liquid-drop model of the nucleus, gives results which are in qualitative agreement with the experimental results. We mentioned on page 302 that the nuclear fission process can also be understood on the basis of this model. If the constants A, B, and C are determined from the experimental material, we get the curve shown in Fig. 152, which follows closely the empirical binding energies.

A very clear survey of the stability of all known nuclei, as determined from experimental mass defects, is shown in Fig. 153. It is actually a three-dimensional representation in which the number of protons is plotted as the abscissa and the number of neutrons as the ordinate, whereas the binding energy of each nucleus (in 1/1,000 mass units = 0.93 Mev) is plotted perpendicularly to the plane of the paper. The resulting surface corresponds to a valley which falls off from the lower left to the upper right corner of the figure, since binding energies are plotted negatively. In this figure, isotopic nuclei lie in planes parallel to the NE plane. Nuclear isobars (nuclei of equal masses), which can change back and forth into each other by emitting electrons or positrons, lie on planes which intersect the NE plane and the ZE plane at 45°. The heavy lines with added numbers drawn in Fig. 153 represent the surface for a particular binding energy.

We must now consider more exactly the stability of different nuclei. Theoretically it follows from the Pauli principle (see page 169) that two protons or neutrons with antiparallel spins can occupy one nuclear energy level, and from this it follows, as in the case of electron shells, that nuclei with fully occupied (closed) energy states, i.e., nuclei with an even number of neutrons *and* protons, must be exceptionally stable. Experience confirms this conclusion. *The α-particle with two neutrons and two protons is the most stable of all known nuclei, and nuclei built up of even numbers of neutrons and protons exceed all others in stability and abundance.* Nuclei with an even number of protons and an odd number of neutrons, or vice versa, are theoretically less stable since only *one* state is fully occupied, saturated, whereas nuclei with an odd number of both protons and neutrons are much less stable. Again experience is in agreement with the theoretical predictions. *The stability and abundance of the even-odd nuclei is considerably smaller than that of the doubly even nuclei, whereas all doubly odd nuclei with atomic numbers over 8 are unstable, i.e., they are β-active.* We shall explain the exceptions, the doubly odd nuclei $_1H^2$, $_3Li^6$, $_5B^{10}$, and $_7N^{14}$ below.

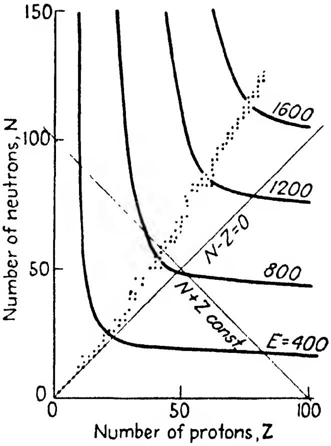


FIG. 153. Dependence of the nuclear binding energy (in 1/1,000 mass units = 0.931 Mev) on the number of protons Z and neutrons N in the nucleus (so-called stability surface). Curves of equal binding energy are indicated.

We come to a more pictorial presentation of this stability problem if we cut isobaric planes through the stability surface on Fig. 153, i.e., planes which cut the Z and N axes at 45° and are parallel to the E axis. We then have cross sections (Figs. 154 and 155) through the valley, and all nuclear isobars having the $(Z + N)$ value under consideration must lie on the curve of intersection of the plane with the surface of the valley. The most stable nuclei lie on the bottom of the valley (i.e., in the minimum of our curves); on the left are those with an excess of protons and on the right those with an excess of neutrons. Those on the slope on the left of the bottom of the valley therefore transform to stable nuclei by emitting *positrons*, those on the right by emitting *electrons*. Actually we find empirically that all nuclei on the left are β^+-emitters (and K radiators, see page 282) and those on the right are β^--emitters. Moreover, according to the theory, only *one* electron or positron can be emitted at a time, never two simultaneously.

Now we consider an isobaric section belonging to the odd mass number 91. It is shown in Fig. 154, where decreasing binding energy is plotted upward on the ordinate, and the neutron excess, $N - Z$, is plotted as abscissa. Here only the lowest nucleus, $_{40}Zr^{91}$, should be stable. Actually there is, according to Mattauch's first isobar law, only one stable nucleus for each odd mass number, whereas all nuclei lying to the left of it transform to the stable nucleus by β^+- or K emission, and all those lying to the right by β^--emission. The situation is more complicated if we cut a section through an even $A = Z + N$ value, e.g., in Fig. 155 through $A = 92$. Both doubly even and doubly odd nuclei belong to this mass number. Since the latter are less stable than the former, the odd-odd curve must lie above the even-even curve; we have here two layers on the bottom of the valley on which the doubly-even and doubly-odd nuclei are situated. Again the most stable nucleus is that with the lowest energy, $_{40}Zr^{92}$. In addition, however, here also the $_{42}Mo^{92}$ nucleus is stable, because it could go over to $_{40}Zr^{92}$ only by a double positron emission, which is forbidden, and it cannot go over into $_{41}Nb^{92}$ because that would require an expenditure of energy. Thus we see that there should be, in general, no stable doubly odd nuclei, but that these should all be β-active. On the other hand, there can be several stable doubly even nuclei (second isobar law of Mattauch's), but they always differ by two units of nuclear charge. The stability of the four lightest doubly odd nuclei $_1H^2$, $_3Li^6$, $_5B^{10}$, and $_7N^{14}$ depends on the fact that the beginning of the valley of Fig. 153 is cut so deep, and thus the slopes are so steep, that the minimum of the upper odd-odd

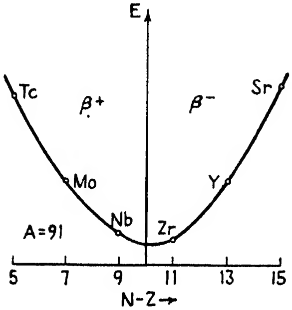


FIG. 154. Isobaric section through the stability surface at an odd mass number ($A = Z + N = 91$).

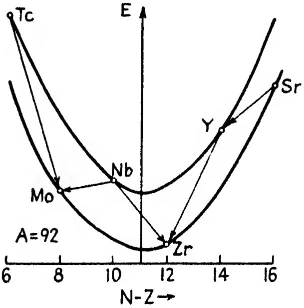


FIG. 155. Isobaric section through the stability surface at an even mass number ($A = Z + N = 92$).

curve with the four nuclei in question is situated *lower* than the nuclei lying somewhat to the side on the even-even curve.

In concluding, we consider briefly the distribution of isotopes, i.e., the number of isotopes with even or odd atomic numbers. From the empirical results Aston derived his famous *isotope rule according to which elements of odd atomic number have at most two stable isotopes, whereas elements with even atomic number usually have more than two.* This may be seen from Table 3 (page 53). By plotting the isobaric sections according to Figs. 154 and 155 it can be shown that this isotope rule is a direct result of Mattauch's isobar laws which we have been able to understand theoretically.

Thus we can predict for any arbitrary nucleus, from its position with respect to the bottom of the "valley of stability," whether it is stable, β^+-unstable, or β^--unstable. In this way it can be predicted (and the prediction has been confirmed) that the new elements belonging to the gaps in the periodic table (see page 309) are not stable, but radioactive.

We have mentioned in Sec. 5-9 that considerable information has been collected recently on nuclear energy levels which now is subject to theoretical analysis. In the early thirties theoretical attempts had already been made to assign to the nucleus a shell structure analogous to that of the atomic electrons (Table 10, page 178). These attempts have been revived by the discovery of the so-called magic-number nuclei. It has been pointed out, namely, that those nuclei which excel their neighbors in stability and abundance are characterized by proton or neutron numbers 2, 8, 10, 14(?), 20, 28(?), 50, 82, and 126. This special stability for certain numbers of protons and neutrons immediately calls to mind the great stability of the closed electron shells of the noble gases (page 169). Consequently, there is general agreement that these magic numbers have something to do with closed nucleon shells. Several theoretical attempts were made very recently to explain them by a construction principle analogous to that of the electron shells which led to the explanation of the periodic table of the elements. In detail, the assignment of orbital momenta to the successive nucleons differs somewhat in the attempts of different authors. This is not surprising, since the interaction between the nucleons is so much larger than that of the shell electrons that the individual-particle model is a much poorer approximation in nuclear than in atomic physics proper. Nevertheless, if we assign quantum numbers to the individual nucleons, we arrive at a simple explanation of the magic numbers. One of several possibilities here is to assume that with some exceptions for every principal quantum number the largest possible orbital quantum number (page 130) leads to the most stable configuration (energy state). This construction principle is different from the construction principle for the electron shells, as Table 10 shows. On this basis the configuration of the closed nucleon shells (and the explanation of the "magic numbers") would be the following:

$(1s)^2$	2
$(1s)^2\ (2p)^6$	8
$(1s)^2\ (2p)^6\ (2s)^2$	10
$(1s)^2\ (2p)^6\ (3d)^{10}\ (2s)^2$	20
$(1s)^2\ (2p)^6\ (3d)^{10}\ (4f)^{14}\ (5g)^{18}$	50
$(1s)^2\ (2p)^6\ (3d)^{10}\ (4f)^{14}\ (5g)^{18}\ (6h)^{22}\ (4d)^{10}$	82
$(1s)^2\ (2p)^6\ (3d)^{10}\ (4f)^{14}\ (5g)^{18}\ (6h)^{22}\ (7i)^{26}\ (4d)^{10}\ (6g)^{18}$	126

The last two orbits might be replaced by the configurations $(4d)^{10}$ $(5f)^{14}\ (2s)^2\ (3s)^2$ or $(5f)^{14}\ (3p)^6\ (4p)^6\ (2s)^2$. The maximum number of nucleons in the different orbits s, p, d, f, g, h, i follows, as on page 170, Table 8, from the Pauli exclusion principle. There is, consequently, a close analogy between electron shells and nucleon shells in this model which provides new evidence for the close interrelation between the different fields of atomic physics.

5-16. High-energy Collision Processes in Cosmic Rays

The processes in and near the nucleus which we have discussed so far were initiated by collisions in which the colliding particle had energies between 10^6 and 10^7 ev. Entirely new and fundamentally significant information about transformations of the elementary particles and the interaction of matter and radiation can be obtained from experiments on radiation coming from outer space, the so-called *cosmic rays*. From measurements of their deviation by a magnetic field in a cloud chamber it has been shown that *the primary cosmic-ray particles have energies of at least* 10^{13} *ev and that it is highly probable that some have energies up to* 10^{16} *ev.*

Only recently by coordination of observation, experiment, and theory has the abundance of complicated primary and secondary cosmic-ray phenomena been analyzed at least to some extent, so that we can begin to separate the primary from the secondary processes. We cannot go into any details of the results pertaining directly to cosmic rays themselves. Moreover, we shall discuss only processes occurring between elementary particles and nuclei in collisions of highest energy. By "highest energy" we understand particle energies well above 10^8 ev. In the most extreme cases the collision energies of the particles exceed even their rest energies, $E_0 = m_0c^2$! What we shall discuss here thus are really the most extreme physical processes that we know.

Only in recent years has conclusive evidence been obtained about the primary particles which fall into the earth's atmosphere from outer space and thus give rise to the complicated abundance of cosmic-ray phenomena. It is generally assumed today that the main constituents of the "primary component" of cosmic rays are high-energy protons. Latest observations, however, leave no doubt that α-particles and, to a smaller extent, also heavier nuclei with atomic numbers up to at least 30 occur as primary particles. It is interesting to note that the ratio of protons to α-particles seems to be 4:1 which agrees approximately with the hydrogen-helium ratio in the sun and other astronomical objects. It seems not impossible, therefore, that *the composition of the primary component of cosmic rays is the same as that of interstellar matter and is determined by the relative abundance of the elements in the universe.*

The origin of cosmic rays, i.e., the source of the tremendous energies of the primary particles, has not yet been satisfactorily explained. Until recently the prevailing opinion was that accelerations of the magnitude encountered with the primary cosmic-ray particles were not possible in our universe as we know it today, but that we had to regard the primary particles as remnants of an earlier stage of the universe such as the primary explosion mentioned above. Recently, however, the idea has found increasing support that this acceleration is still possible due to vortical magnetic fields, comparable to the acceleration in the betatron. Of great interest in this respect is Babcock's recent discovery that some fixed stars have a magnetic field which varies in amplitude and sign and thus, with exception of the missing guiding field (page 261), are similar to a huge betatron. Vortical magnetic fields, moreover, do exist also in connection with the solar spots and flares of our sun. Some physicists, for this and other reasons, regard cosmic rays as a phenomenon of our solar system, and actually the cosmic-ray intensity shows some remarkable changes which coincide with prominent solar flares. It is not improbable, therefore, that there is more than one source of cosmic rays.

The great difficulties encountered in analyzing cosmic rays can be understood if we think of the large number of interacting processes which occur with particles of 10^{10} to 10^{16} ev maximum energy, such as explosions and cascade showers, meson decay and other meson processes, bremsstrahlung, pair production, ionization, etc., all to be discussed in detail below. Even at the greatest heights which have been studied extensively by recording instruments, about 30 km, we still do not encounter exclusively primary radiation, but a rather complex radiation from which it is not too simple to draw conclusions about the primary particles. It is therefore of special interest that now rockets with cosmic ray equipment are fired into heights up to 200 km. They may aid materially in clearing up cosmic-ray phenomena.

We shall limit our discussion to the fundamentally important processes. We distinguish between processes occurring as a result of primary particles and those due to secondary particles. Among the secondary particles produced by the primary cosmic-ray particle we find high-energy photons and electrons of both signs, whose processes we shall discuss in the next section. In addition to these there is another group of secondary particles which are formerly unknown elementary particles, on which the attention of all physicists who are interested in fundamental research is concentrated. These particles are the different types of *mesons*. They will be discussed in the second section following.

First, we shall discuss the processes initiated by the primary particles, i.e., extremely high-energy protons, and to a small extent, neutrons, α-particles, and heavier nuclei. In passing through normal matter, fast nucleons lose kinetic energy by ionization, whereas bremsstrahlung processes of such heavy particles are probably not very important.

The processes initiated in nuclear collisions of fast protons and other primary cosmic particles have been mentioned to some extent on page 286. The colliding particle can transfer energy and momentum in direct collisions to a few nucleons of the nucleus. These nucleons can then be ejected. By a central collision with a heavy nucleus, on the other hand, the whole nucleus can be heated up as a result of the energy acquired from the colliding particle, so that subsequently several nuclear particles can evaporate or the whole nucleus may explode (Fig. 156). According to Schopper this can even occur in steps. In such a nuclear explosion occasionally some primary mesons, to be discussed on page 332, can also be emitted.

Finally, it seems *that in collisions between nuclei and high-energy nucleons an entirely new and fascinating elementary process can occur — the production of a number of electrons and mesons of both signs (and probably nucleons also) by materialization of the kinetic energy of the colliding nucleon.* Heisenberg has expressed the opinion that such a *materialization of kinetic energy* has contributed to the "explosion shower" shown in Fig. 156 and that it thus is not a normal nuclear explosion, such as Fig. 141. Further conclusive evidence has been published meanwhile for such processes.

This process would represent a large-scale parallel to the production of an electron pair, to be discussed in the next section, in a collision of a high-energy photon with a nucleus. That such a materialization process, which of course obeys Einstein's energy equivalence equation (5-6), occurs only in nuclear collisions, depends on the fact that a strong field is necessary to transform kinetic or radiation energy into materialized energy (mass).

Electron-pair production deals with what, in the wave language, is called an interaction between the electromagnetic field (photon) and an electron-wave field, and this interaction can be described and understood on the basis of quantum mechanics. However, quantum mechanics fails in the quantitative description of the nuclear field, which corresponds to the π meson (page 332) as the electromagnetic field corresponds to the photon. Therefore, we cannot expect an explanation of multiple particle production from present-day quantum mechanics. An attempt of Heisenberg's to develop quantum mechanics so as to include nuclear phenomena was mentioned on page 241. This new rough theory also accounts for the multiple production processes, the production of several elementary particles in a single elementary process. Because there is

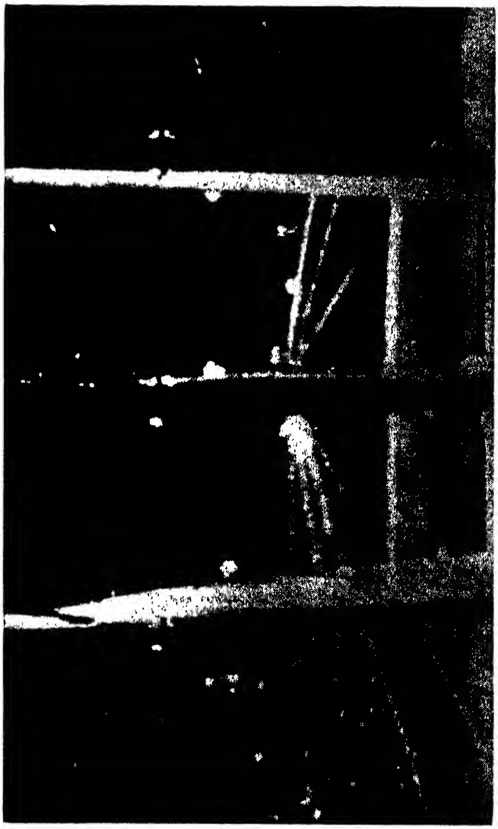


FIG. 156. Cloud-chamber photograph of an explosion shower. (*After Fussell.*)

an intimate relation between these materialization processes and the latest development of the atomic theory, physicists are looking forward eagerly to learning more about these collision processes of highest energy.

5-17. Collision Processes of High-energy Electrons and Photons. Electron Pairs and Dirac's Hole Theory. Cascade Showers

As secondary products of primary cosmic-ray collisions and of the meson decay (which will be discussed in the next section, page 332), high-energy electrons and photons play a considerable role in the secondary cosmic-ray phenomena. Investigations of these electrons with the cloud chamber and photographic plate led to the discovery of a number of theoretically important elementary and secondary processes, which will be discussed in the following.

A high-energy electron, in penetrating matter, can interact with the electron shells of atoms, and by ionizing collisions dissipate its kinetic energy and thus slowly come to rest. On the other hand, it can, in colliding with a nucleus, be stopped by the strong electric field of the nucleus and thus produce extremely short-wave bremsstrahlung (page 110), i.e., γ-rays with $\lambda \ll 0.1$ A. The ionization becomes the less probable, the greater the kinetic energy of the particle, because, if the electron passes through the electron shell too fast, there will not be enough time for a transfer of energy to shell electrons. Thus the faster an electron is, the less it is hindered by the electron shells, and finally it can be stopped only by colliding with nuclei under emission of bremsstrahlung. Exact computations confirm this: *For electron energies above* 10^8 *ev, the energy loss due to bremsstrahlung begins to exceed that due to ionization.* It is of interest that this energy region now has come within the experimental range due to the betatron and electron synchrotron, as mentioned on page 260.

We can describe the bremsstrahlung process by saying: the kinetic energy of an electron is expended to produce a bremsstrahlung quantum, i.e., a very high-energy photon, a γ-quantum. The directly reversed process, the absorption of a high-energy photon by an atomic electron (see page 107) that converts the total energy of the photon to kinetic energy of an electron, is rare. Nevertheless, it must be considered in addition to the Compton effect, page 194, where in a collision the photon imparts a fraction of its energy to an electron.

As soon as the energy of a photon exceeds 1×10^6 ev, which corresponds to twice the rest mass of the electron, a new process is possible. *A high-energy photon, by losing its entire energy in the field of a nucleus, "produces" an electron pair, an electron and a positron* (Fig. 157). The fact that only an electron *pair* can be produced for each photon that disappears follows from the charge balance: an uncharged particle such as the photon cannot produce a charged particle; it can only produce a pair of particles of opposite sign whose electric charges compensate each other. That this pair production can take place only in the vicinity of a nucleus, i.e., during a collision of the photon with a nucleus, follows, first, from the laws of conservation of energy and momentum according to which the nucleus must take over part of the momentum, and, second, from the fact that the electrostatic field of the nucleus has to act as an agent between the vanishing photon and the electron pair. *The definite proof of the process of pair production, established by means of the cloud chamber (Fig. 157), must be regarded as one of the most important results of modern physics, since here for the first time the direct conversion of energy into matter, according to the equation*

$$E = mc^2 \qquad (5\text{-}6)$$

is observed. Also the reverse process has been found, the "annihilation" of an electron-positron pair upon production of two γ-quanta, again in agreement with the complete equation

$$E = m_0c^2 + \frac{m}{2}v^2 \qquad (5\text{-}37)$$

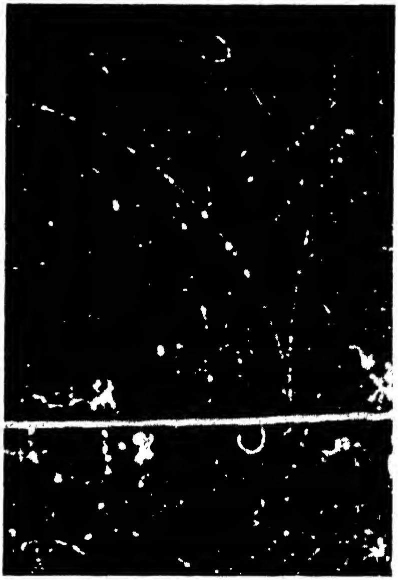


FIG. 157. Cloud-chamber photograph of the production of two electron pairs by high-energy photons (hard γ-rays of 17.6 Mev). Photographed in a magnetic field of 2,500 gauss. (*By Fowler and Lauritsen.*)

DuMond recently succeeded in measuring directly, with his two-meter bent-crystal X-ray spectrograph, the wavelength of the two γ-quanta emitted upon annihilation of an electron pair. He found it to be 0.02427 A, which is in excellent agreement with the theory according to (5-37) and (3-10).

The discovery of positrons and pair production, as well as their annihilation, was echoed resoundingly by theoretical physicists, because Dirac in his relativistic theory of the electron had obtained results which seemed to be confirmed by these discoveries. The fundamental ideas of this much discussed *hole theory of Dirac* can be outlined briefly as follows:

If all possible energy states of the "ordinary" negative electron are plotted in a term diagram, including the rest energy of the electron corresponding to its rest mass, $E_0 = m_{e0}c^2 = 0.5$ Mev, then the energy of a free electron at rest lies 0.5 Mev above the zero line (Fig. 158). The discrete states of the electron, when bound to the atom, lie somewhat lower and the kinetic-energy states of the continuum of the free electron lie higher. Now it follows from Dirac's theory that, in addition to these known energy states, there must exist energy states which correspond to negative masses and therefore, according to Fig. 158, must lie more than 0.5 Mev *below* the zero-energy line. The obvious objection that the existence of these negative states should make itself evident by the transition of electrons from the known ones to these new states (and thus the H atom could not be stable), was met by Dirac with the assumption that normally these states are filled and, according to the Pauli principle, transitions into them are impossible. In order that an electron from one of these (unobserved) states be raised into the "upper world," energy in the amount of 1 Mev must be expended; for instance, *a correspondingly high-energy photon must be absorbed. Then the result is the simultaneous "production" of a negative electron and a positively charged "hole" in the negative energy states. It can be shown that such a hole would have the properties of a positron, and the sketched absorption process would mean the production of an electron pair. The disappearance of an electron pair by transformation*

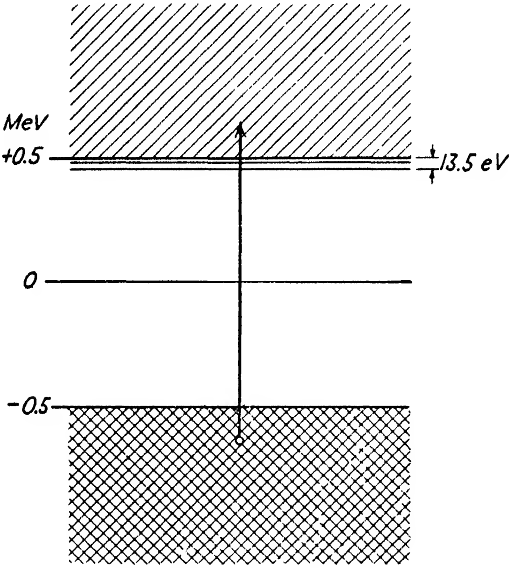


FIG. 158. Complete energy-level diagram of an electron according to Dirac's relativistic theory of the electron. Arrow shows "production" of a positive and a negative electron by absorption of a photon with an energy above 1.1 Mev (process of Fig. 157).

into radiation energy means, in this model, simply that there is a transition of an electron to a hole in the continuum of the negative energy states. That positrons are observed relatively rarely can be explained directly by the hole theory: *since our observation space is more or less densely filled with matter, and thus with normal electrons, there is always an electron in the neighborhood of a positron which can slip into the hole and thus cause the disappearance of the positron.*

Thus Dirac's hole theory, which is only very sketchily presented here, can explain satisfactorily the positron and certain other phenomena (such as Klein's paradox) which are not discussed here. Nevertheless it cannot be denied that the assumption of an energy continuum corresponding to negative mass values, which is normally filled with electrons and thus cannot be observed, is a very unsatisfactory feature of this relativistic theory of the electron and positron.

FIG. 159. Cloud-chamber photograph of five cascade showers of electrons and positrons, produced in a lead plate by parallel incident cosmic-ray particles. (*After Brode and Starr.*)

In cosmic-ray observations by means of the cloud chamber, so-called "showers" of positive and negative electrons are often observed. Figure 159 shows some examples of showers produced by high-energy electrons or photons in a lead plate. If, instead of one thick plate, a series of plates is arranged in the cloud chamber (Fig. 160), it can be seen that these showers are not produced in one elementary process like that of Fig. 156, but by a sequence of processes, a cascade. For this reason they are called *cascade or multiplication showers.* They are caused by a *rapid succession of the two elementary processes discussed above, bremsstrahlung and electron-pair production.*

If a very fast, i.e., high-energy, electron falls on a lead plate, it produces, in colliding with a nucleus, a correspondingly very high-energy bremsstrahlung quantum (photon). After a short flight this photon produces an electron pair, whereby each of these two particles takes half the energy of the photon with it. Each electron and positron, after a short time, again produces a bremsstrahlung photon, these two photons produce two electron pairs, etc. Thus, as an end result, the energy of the primary electron will be distributed among a large number of electrons and positrons, as a consequence of this multiplication process. For example, a primary electron with an energy of 10^{11} ev can produce about 1,000 particles in a multiplication shower in a 5-cm layer of lead. Here lead with its high nuclear charge and density is used to study multiplication processes because it requires little space so that the whole process can take place and be observed in a small cloud chamber. Multiplication showers also take place in our atmosphere, but the cascading extends over a much thicker layer. In conclusion, it should be emphasized that a multiplication shower can be initiated by a very high-energy photon just as well as by a fast electron.

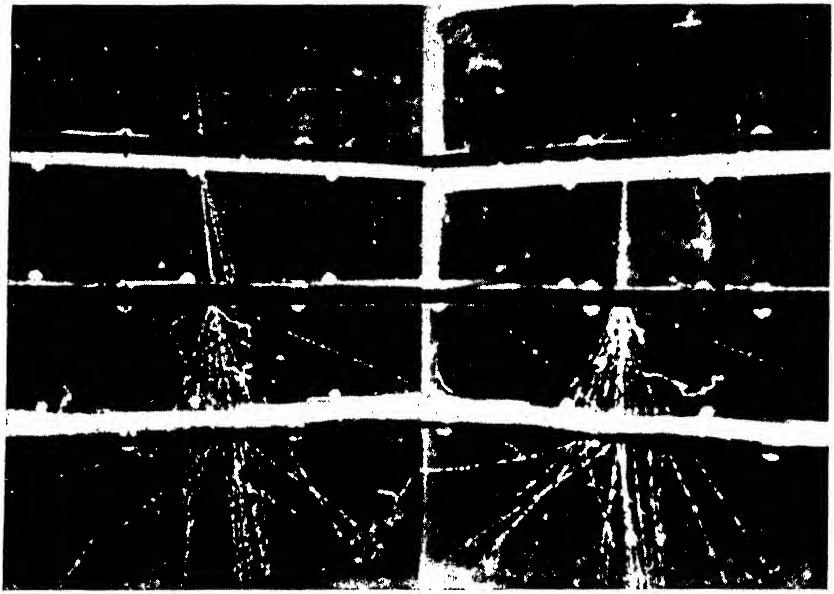


FIG. 160. Cloud-chamber photograph of a cascade shower originating from numerous consequent lead shields, providing evidence for the multiplication mechanism of cascade showers. (*After Fussell.*)

According to Blau and Wambacher (1937), extremely high-energy electrons or photons can also cause the nuclear explosions mentioned on pages 286 and 324, in which a nucleus is heated so high by the great energy acquired in a collision that a large number of nucleons evaporate from it.

5-18. Mesons and Their Role in Cosmic-ray Physics

An entirely new elementary particle was discovered in cosmic rays in 1936 by Anderson and Neddermeyer, the *heavy electron or meson.* On cloud-chamber photographs, taken in a magnetic field, tracks of a particle were occasionally observed whose ionization (droplet density) was too small for a proton but greater than that expected for an electron of such high velocity, as deduced from the small curvature of its path in the magnetic field. Further investigation produced the unequivocal result that this was an elementary particle hitherto unknown, whose charge was the same as that of the electron, but whose mass was between the electron and the proton. For this reason it is called the *meson*, or, occasionally, the mesotron. The discovery of this particle was not entirely unexpected, since in 1935 Yukawa had postulated the existence of such a charged elementary particle with a mass between that of electron and proton, in his attempt toward a new theory of nuclear forces.

He set up a wave equation which should represent the nuclear forces in the same way in which electric forces are described by an electromagnetic wave equation. This new wave equation contains a term which accounts for the short range of the nuclear forces. In the particle language, this particular wave field corresponds to the meson, whose existence was inferred by Yukawa from such arguments. We shall return to Yukawa's theory on page 337.

However, the rapid development of meson research, especially in the last few years, has shown clearly that the meson problem is much more complicated than it seemed at first. The feeling prevails today that *the meson problem contains the key for understanding many yet unexplained and difficult problems of nuclear physics in the broader sense, and perhaps even for the whole theory of elementary particles*, to be discussed shortly.

Although the situation is by no means clear today and new experimental results and theories are published almost daily, we believe that *the existence of various kinds of positive and negative mesons and at least one neutral meson, often called the neutretto, has been definitely proved.* Furthermore, it is certain that the cosmic-ray meson discovered by Anderson is *not* identical with the Yukawa particle which is supposed to be responsible for nuclear binding. While the mass of the Anderson cosmic-ray meson (called the *μ meson*) is 212 ± 5 electron masses, Powell and Occhialini have recently discovered the π meson with a mass of approximately 280 m_e. This new meson was found during cosmic-ray studies of photographically recorded particle tracks, among the fragments from nuclear explosions (page 324). Meanwhile this observation has been confirmed by experiments with the Berkeley cyclotron where these *heavy mesons or π mesons* could be obtained by bombarding light nuclei with 380-Mev α-particles. At the present time it seems by no means certain that the different types of mesons are exhausted with these two kinds, each of which occurs positively and negatively, and with one *neutral meson* (*π^0 meson or neutretto*) with a mass of around 270 m_e, whose existence seems certain. Leprince-Ringuet has found evidence for mesons with masses between 700 and 1,000 electron masses, and more than half a dozen similar observations leave little doubt as to the existence of at least one *very heavy meson or τ meson.* The existence of lighter mesons down to 4 m_e, which has been claimed by some observers, however, is very doubtful.

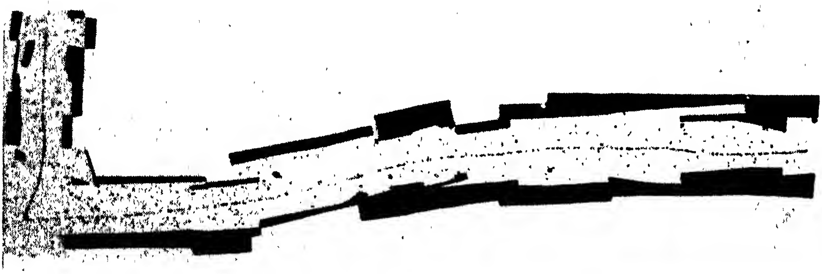


FIG. 161. Photographic track of a π meson, incident vertically, which decays into a μ meson (progressing to the right) and a nonionizing particle, probably a neutrino. The figure is a mosaic made of several microphotographs. (*Courtesy of C. F. Powell and G. P. S. Occhialini.*)

Both types of mesons about which we have some detailed knowledge are unstable. The half-life of the π meson is approximately 1×10^{-8} sec, that of the μ meson 2.2×10^{-6} sec. Moreover, we know that the π meson of 280 m_e decays into a μ meson of 210 m_e and, most probably, a neutrino and an energy of approximately 35 Mev. (Fig. 161). The μ meson, during its decay, certainly produces electrons with a continuous energy spectrum, so that here the situation seems to be analogous to that of the radioactive β-decay (page 279). The decay of the μ meson must be a triple decay just as that of a β-active nucleus, which decays into the new nucleus, an electron, and a neutrino. It seems certain that the μ meson decays into one electron and two neutrinos (Fig. 162). The fact that decay electrons with energies up to 50 Mev have been found is good evidence for this process, since the total available energy of 210 m_e equivalent to 100 Mev has to account for the masses and kinetic energies of *all* the decay products.

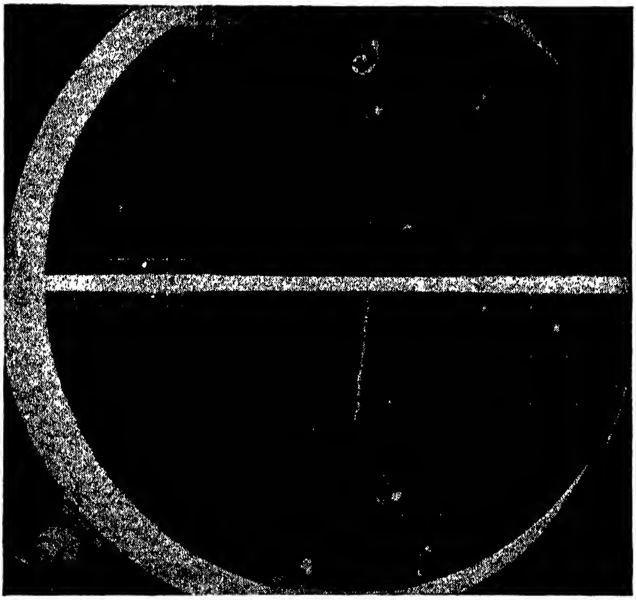


FIG. 162. Cloud-chamber photograph of the decay of a μ meson. The fast μ meson is decelerated in the lead plate and therefore ionizes heavily below the plate until it decays into an electron and, most probably, two neutrinos. The electron track extends from the end of the meson track toward the lower right corner. (*Courtesy of R. W. Thompson.*)

Thus there is no doubt that the μ meson of 210 m_e mass, which forms a very essential part of the cosmic rays penetrating to the lower layers of our atmosphere, has to be regarded, just as many light particles, as a secondary particle produced by decaying mesons. This accounts for the experimental result, at first rather disturbing, that the μ meson does not interact greatly with the nucleus and thus is not emitted by it. Moreover, for the same reason, it cannot be identified with the Yukawa particle. On the other hand, Powell and Occhialini were able to show that π mesons *are* emitted, together with nucleons, in nuclear explosions. The π meson also interacts with nuclei after it has been emitted, i.e., it can induce the ejection of several nucleons from a nucleus. There is, consequently, much evidence in favor of identifying the π meson with the Yukawa particle, moreover as theoreticians had already expected a life-time of 10^{-8} sec for the "nuclear-force meson." The positive π meson then would be produced and emitted if, upon participation of nuclear kinetic energy, a proton in the nucleus is transformed into a neutron, whereas the negative π meson is emitted during the reverse process. We shall return to the close relation of this process to the binding of the nucleons on page 337.

The life of the presently well-known mesons seems to be the following. Primary positive and negative (and most probably also neutral) π mesons are produced in nuclear explosions, which are excited in nature by the primary cosmic-ray protons, and in the laboratory by highly accelerated α-particles (and in the future certainly also by protons and electrons). In passing through matter, the negative π mesons are captured, in part, within their life-time of 10^{-8} sec, by the positively charged nuclei, and produce here additional nuclear reactions such as in Fig. 141. The positive π mesons, because of the repulsive electric forces, cannot be captured by the nuclei and, just as the negative π mesons which are not captured within 10^{-8} sec, decay after their life-time into the lighter (positive or negative) μ mesons and one neutrino each. The neutral π^0 mesons decay, after a very short life-time of 10^{-14} seconds, into two photons with the combined energy of 135 Mev. An appreciable fraction of the negative μ mesons is also captured by nuclei, because their small interaction with nuclei is compensated by their half-life of 2×10^{-6} sec which is 200 times that of the π meson. The apparent half-life of the negative μ meson, for this reason, is smaller than that of the positive μ meson, and it depends on the properties of the matter through which they move. The uncaptured negative μ mesons, together with the positive μ mesons, decay, as discussed above, most probably into electrons and neutrinos (Fig. 162).

There is nothing definite known about the spin of the different mesons. The π meson has to have an integral spin, most probably the spin zero. The μ meson, because of its decay into three particles, might have a half-integral spin and probably has the spin $\frac{1}{2}(h/2\pi)$, at least if it definitely decays into an electron and two neutrinos. This is about all we know of the meson problem, the further development of which is awaited with excitement by all atomic physicists.

In passing through matter, mesons lose part of their energy, just as do electrons, in ionizing the electron shells of atoms. Bremsstrahlung from nuclear collisions is less important for mesons than for electrons, because the intensity (probability) of the retardation process is inversely proportional to the mass of the colliding particles. However, the meson is frequently scattered by the nucleus (analogous to the Compton effect). With electrons, this process is important only for energies below 10^6 ev, whereas for larger energies bremsstrahlung becomes more important. Whether high-energy mesons, just as protons (page 324), can materialize their kinetic energy in the form of several particles (electrons, mesons, etc.) by multiple processes, is not yet known. Research with artificial high-energy particles from synchrotrons, bevatrons, etc. (page 263), will, no doubt, be of utmost importance in supplementing cosmic-ray results concerning the meson problem.

5-19. Nuclear Forces and the Problem of Elementary Particles

Two points in our present treatment of nuclear physics are most unsatisfactory. In the first place, we have not made any definite statements about the nature of the nuclear forces which bind protons and neutrons within the nucleus. This problem of the nuclear forces, which is a difficult one and has by no means been solved satisfactorily as yet, has to be treated in somewhat more detail. In the second place, we have encountered, during our treatment of atomic physics, quite a number of so-called "elementary particles," and we have to discuss what is known at present about their number, their properties, and their role from a theoretical point of view.

With respect to the nuclear forces we want to emphasize that they apparently *cannot* be understood from the standpoint of classical physics, because none of the nuclear processes which depend intimately on the nuclear forces has a classical explanation. The transformation of neutrons into protons, and vice versa, in the nucleus under emission and absorption of positive and negative π mesons, the processes of electron-pair production, meson production, and multiple processes, which have been discussed in the last sections, evidently lie far outside the realm of classical physics. We therefore have to consider the problem of nuclear forces from the quantum-mechanical point of view, according to which corpuscles and wave field are but two phenomena of the same physical reality. Let us consider from this point of view, at first, the forces effecting the binding between the nucleus and the electron shells of an atom.

In the standard description, the electrostatic field between the positive nucleus and the negative electron is responsible for the stability of the atoms. An excited atom, in terms of this theory, "contains" a larger amount of electrostatic field energy, and this excess electrostatic energy is emitted as a photon (or an electromagnetic wave, respectively), when the atom returns to the normal energy state. As the electrostatic field is but a special case of the electromagnetic field, we can make two statements about the binding between nucleus and electron shells which amount to the same: "The binding is effected by the electromagnetic field," or "The binding is due to the possibility of emission and absorption of photons by the atom." Dirac has developed this idea into a rigorous and concise radiation theory in his quantum electrodynamics.

Now we know that the protons and neutrons in a nucleus can transform into each other by emission and absorption of mesons (or in the older theory of electrons, positrons, and neutrinos), and that these mesons (or the secondary electrons and positrons) are actually emitted by β-decaying nuclei. It, therefore, *seems reasonable to attribute the mesons to the field of the nuclear forces in a similar way as the photons have to be attributed to the electromagnetic field.* During our discussion of the β-decay on page 281 we have already indicated the principal ideas of Fermi's theory of β-decay, based on the analogy between photon emission by atoms and electron-neutrino emission by nuclei. This theory, however, led to a binding energy of the nuclei which was far too small. Yukawa therefore replaced it by the hypothesis that primarily the exchange of what we today call π mesons effects the binding, and that in the case of β-decay these π mesons finally decay into electrons and neutrinos. Because of the mutual electrostatic repulsion of the protons, however, an electric, or, more generally, an electromagnetic, field also contributes to the nuclear forces. In the corpuscular description, this electric nuclear field is equivalent to the possibility of emission and absorption of photons which, because of the large nuclear energy, in this case must be γ-quanta, and we know already that γ-ray photons are actually emitted and absorbed by nuclei. We thus have the following coordination of fields and corpuscles, which is the basis for the understanding of atomic and nuclear forces:

	Atom shells	Atomic nucleus	
Constituent particles....	Electrons	Protons and neutrons	
Force field.............	Electric field	Electric field + nuclear field	
Particles emitted upon energy changes.......	Photons	Photons (γ-quanta)	π mesons ↓ Finally electrons, positrons, and neutrinos

The nuclear forces proper (after separation of the electric contribution) thus depend on an exchange of mesons (with positive or negative charge). The proton transforms into a nuclear neutron by absorption of a negative (or emission of a positive) meson; the neutron upon emission of a negative (or absorption of a positive) meson transforms into a proton. The energy balance is kept in order because the processes occur within the nucleus, where energy is available. This continuous transformation of neutrons into protons and the reverse can also be regarded as an exchange of these two particles. In this description of the nuclear forces as exchange forces, we recognize the relation to the quantum-mechanical exchange resonance, discussed on page 230, and to the binding of atoms in homopolar molecules. In a pictorial description (page 357) these binding forces are due to the fact that electrons encircle both nuclei or, in other words, that they are exchanged between the two atoms.

Conclusive experimental evidence for the correctness of this exchange theory has recently (1947) been found in Berkeley. Scattering of 90-Mev neutrons by nuclei resulted in a much larger number of protons scattered in the direction of the colliding neutrons than could be understood theoretically under consideration of the binding energy of the scattering nuclear protons. The only possible explanation of this proton excess seems to be that *a certain fraction of the primary neutrons has been transformed, during the collision, into protons by meson exchange with nuclear protons.*

It is obvious that this theory of nuclear forces, which had been developed by Yukawa before the discovery of the mesons, has received strong support by the discovery of a particle (π meson) which agrees in essential properties with those expected for Yukawa's particle. We therefore can describe the situation in the wave and particle picture in the following way:

Wave representation: neutron (proton) produces nuclear field; nuclear field interacts with proton (neutron).

Particle representation: neutron (proton) emits negative (positive) meson; meson is absorbed by proton (neutron).

We want to indicate by this discussion that the nuclear forces can be treated by the quantum-electrodynamical theory, though a satisfactory quantitative treatment at the present time is still missing, and certainly will be more complicated than Yukawa's first attempt. We must await the results of the present efforts of a large number of theoretical physicists, including men like Dirac, Heisenberg, Oppenheimer, and Schwinger.

From the meson theory of nuclear forces there also follows the explanation for the fact that the neutron has a magnetic moment (though it has no electric *charge!) and that the magnetic moments of proton and neutron, according to page 269, are different from the value of the unit of the nuclear moment,* as expected from the primitive theory of atomic magnetism. In quantum mechanics the "possibility to undergo a certain process" is a definite part of the properties of an atomic system. Since protons and neutrons both have the ability to emit a positive or negative meson, respectively, their magnetic moments are not those to be expected for the "proton alone" or the "neutron alone"; the latter should be zero according to classical physics. These nucleons have, moreover, an additional nuclear moment due to the meson which, with a certain probability, can be emitted by them. From Eq. (3-84) we see that the contribution of this "virtual meson" to the magnetic moment is comparatively large because of its comparatively small mass. The magnetic moment of the neutron thus stems entirely from its ability to decay into a proton and (finally) an electron. The direction of this moment is opposite to that of the mechanical neutron spin and, therefore, in a first approximation, is negative by the same amount as the meson contribution to the proton magnetic moment is positive.

This last discussion again was evidence for the strange properties of the short-lived elementary particles, mesons, positrons, and neutrinos, and for the peculiar role which they play in atomic physics. In concluding this section, we therefore want to review briefly the statements which we can make at present about the *elementary particles.* As such we designate particles *which occur in elementary atomic processes and, according to our present knowledge, are not composite particles in the normal sense* (such as is the α-particle).

Besides the three elementary particles which form our whole material world, namely protons, neutrons, and electrons, we know of a number of elementary particles which, with exception of the last one, appear and disappear only in nuclear processes; these are the mesons, the positron, the neutrino (and antineutrino, page 283), and the photon. All these elementary particles are characterized by four main properties—their charge, their rest mass, their mechanical spin, and their magnetic moment. For the unstable particles, the mean life and the decay products are additional characteristic properties. However, those elementary particles which are stable when in an isolated state (vacuum), can also undergo characteristic changes by interaction with other particles. The pair production and the annihilation of electron and positron upon absorption and emission of photons, respectively, and all the other collision processes which have been discussed in Secs. 5-16 to 5-18, belong to the essential properties of these elementary particles. Table 12 presents some of their most important properties:

Table 12. Properties of Elementary Particles

Particle	Symbol	Charge	Rest mass	Spin	Stability in vacuum
Proton . . .	p	$+e$	1,837 m_e	$\frac{1}{2}(h/2\pi)$	Stable
Neutron . . .	n	0	1,839 m_e	$\frac{1}{2}(h/2\pi)$	Unstable
Electron	e	$-e$	m_e	$\frac{1}{2}(h/2\pi)$	Stable
Positron . .	e^+	$+e$	m_e	$\frac{1}{2}(h/2\pi)$	Stable
τ meson	τ	$\pm e$	700–1,000 m_e?	?	Unstable
π meson	π	$\pm e$, 0	280 m_e	0?	Unstable
μ meson	μ	$\pm e$	210 m_e	$\frac{1}{2}(h/2\pi)$?	Unstable
Neutrino . .	ν_0	0	0	$\frac{1}{2}(h/2\pi)$	Stable?
Photon . . .	$h\nu$	0	0	$1(h/2\pi)$	Stable

We do not have any concise theory, or even an idea, of how to explain the fundamental properties of the elementary particles. This is due, at least partly, to the fact that quantum electrodynamics is still not sufficiently developed to account for the properties and interaction of the wave fields which, according to our discussion above, have to be coordinated with the particles. We therefore do not know, for instance, whether our Table 12 is essentially complete, or whether we have to expect the discovery of additional elementary particles, even aside from other mesons which are not improbable at all. There has been some discussion about the existence of a negative proton, and its discovery has been claimed at least once. It has not been confirmed so far, however. In contrast to the neutral mesons (which seem to exist also with masses higher than 70), the existence of a neutral electron is not too probable, at least if the electron mass, according to the derivation of the "classical electron radius" (page 36), is essentially electrostatic mass. The theory of the elementary particles thus is the most important of the major unsolved problems of present-day atomic physics. One of the many difficulties of the problem is the fact that the elementary particles must have a finite size, because point-like particles would imply the much discussed difficulties of an infinite eigenenergy and thus an infinite mass. This is obvious for the Coulomb energy e^2/r which approaches infinity for $r \rightarrow 0$. On the other hand, the elementary particles cannot have extension in space in the normal sense, because that would imply a structure (continuous or discontinuous) which seems irreconcilable with the concept of an elementary particle. As all determinations of the radius of elementary particles (pages 18 and 266) point consistently to values of the order of 10^{-13} cm, March recently has tried to remove the indicated difficulty by the assumption that the accuracy of a single measurement of a length is, by principle, limited to a fundamental length l_0 of the order of 10^{-13} cm. The reason would be that the "yardstick" itself consists of elementary particles, and measuring a length actually means determining the coincidence of two elementary particles. March now asserts that the possibility of resolving two particles is limited by l_0, in the sense that two particles seem to coincide if the distance of their centers is smaller than l_0. This would mean that a particle could, by principle, not be localized more accurately than up to l_0. It can be proved that the application of the uncertainty principle, Eqs. (4-19) and (4-22), under this assumption leads to a limitation of the largest possible exchange of momentum and energy in a single collision process. This would automatically remove all difficulties. The value of these considerations cannot be judged clearly before the role of this length l_0 in the mathematical theory, which must comprise a theory of the elementary particles, is more precisely defined. Heisenberg's attempt in this direction has already been mentioned (page 241). We shall bring up this problem again in the next section.

First, however, we want to indicate the fundamental change which the concept of the elementary particle has undergone, and how far it is removed from classical physics. In classical physics the elementary particles (for instance, about 1900 the atoms) played the role of the invariable "building stones" of our material world, on which forces acted and produced accelerations. Not much that is valid has remained of this classical picture. *The invariability of the elementary particles is lost:* Electron pairs are produced and disappear just as do photons and neutrinos. Mesons are produced and decay. Of the heaviest elementary particles, the neutron undergoes a radioactive decay, whereas the proton, in the nucleus, can transform itself into a neutron by absorption of a negative meson. Moreover, it is not impossible that even these heaviest elementary particles can be produced by materialization of energy in extreme cosmic-ray collisions, though this process has not yet been definitely proved. *Not only the invariable stability of the classical elementary particle has disappeared; also the second characteristic property of the classical theory, the clear distinction between a particle and the force which acts upon it, has no clear meaning any more,* since we know that the elementary particles themselves produce the fields by which they interact with other particles. Moreover, we also know that a distinction between a particle (meson) and its force (nuclear force field) is basically impossible.

An entirely new concept of nature thus stems from atomic physics, and it is understandable if more than a few experimental scientists hesitate to accept this development of modern physics with all its conceptual difficulties and the impossibility of a clear pictorial presentation. However, all these scientists, as well as the doubtful reader, may consider that this development was enforced by the discovery of new natural phenomena (particularly in cosmic-ray physics) which could not be described in terms of the old classical picture of physics. Only the future will reveal, after the present theoretical attempts have been developed into a consistent theory, what we gain by abandoning the classical pictorial theory.

5-20. The Problem of the Fundamental Constants of Physics

Closely related to the problem of the elementary particles is one of the most fascinating problems of modern physics, that of the fundamental constants of nature. Taking into account all our present knowledge of atomic physics, we ask to what ultimate constants and fundamental laws all physics can be reduced, and what we know at present about these fundamental constants.

We know that all atoms, and with them our whole material world, consist of only three elementary particles: protons, neutrons, and electrons. We might even speak of only two by taking proton and neutron with their nearly identical mass as two states of one elementary particle, the nucleon. The nucleon and the electron, then, provide three constants, the nucleon mass M, the electron mass m, and the elementary charge e. We have, furthermore, the universal constants contained in the fundamental laws of physics, Planck's quantum of action h, the velocity of light c, and the gravitational constant f. These six constants thus should be sufficient to develop our whole structure of physics:

$$\left.\begin{aligned} M &= 1.673 \times 10^{-24}\ \text{gm} \\ m &= 9.108 \times 10^{-28}\ \text{gm} \\ e &= 4.803 \times 10^{-10}\ \text{gm}^{\frac{1}{2}}\ \text{cm}^{\frac{3}{2}}\ \text{sec}^{-1} \\ h &= 6.626 \times 10^{-27}\ \text{gm cm}^{2}\ \text{sec}^{-1} \\ c &= 2.998 \times 10^{10}\ \text{cm sec}^{-1} \\ f &= 6.664 \times 10^{-8}\ \text{gm}^{-1}\ \text{cm}^{3}\ \text{sec}^{-2} \end{aligned}\right\} \qquad (5\text{-}38)$$

According to Heisenberg's extension of quantum mechanics, which we mentioned on page 241, there seems to be little doubt that we need, as an additional constant, a fundamental smallest length. Its exact value is not known, but it must have the order of magnitude 10^{-13} cm. The question thus arises whether this new constant is an independent seventh fundamental constant, or whether it can be derived from the constants (5-38). It is interesting to note that from the constants (5-38), even two combinations can be derived which have the dimension of length and the correct order of magnitude. One of them is the classical radius of the electron, known already from page 36,

$$r_e = \frac{e^2}{m_e c^2} = 2.82 \times 10^{-13} \quad \text{cm} \tag{5-39}$$

the other, often forgotten, is the so-called Compton wavelength of the proton

$$l_0 = \frac{h}{M_p c} = 1.32 \times 10^{-13} \quad \text{cm} \tag{5-40}$$

With respect to this last figure, the author has pointed to the fact that it agrees, within the accuracy of the empirical values, with the empirical radius of the proton (page 266). Therefore, it seems reasonable to regard this quantity, in contrast to (5-39), as the natural unit for nuclear dimensions, since it contains, besides the two most fundamental physical constants h and c, only the mass of the most stable nuclear particle, that of the proton, and agrees with its radius. As long as we have no further information on the value of the fundamental length, l_0, it seems not unreasonable to regard (5-40), or maybe its value divided by 2π, as the fundamental smallest length. We certainly shall not be far off with this assumption. This would mean that we have a relation between l_0, M, h, and c which can be written

$$\frac{h}{Mcl_0} \approx 1 \tag{5-41}$$

However, from the six fundamental constants (5-38), three dimensionless constants can be formed, so that, in principle at least, it should be possible to reduce the number of independent fundamental constants of physics to three. In the first place we have the mass relation of nucleon and electron

$$\frac{M}{m} = \beta = 1{,}837 \tag{5-42}$$

In the second place, the dimensionless quantity e^2/hc, which, multiplied by 2π,

$$\frac{2\pi e^2}{hc} = \alpha = \frac{1}{137} \tag{5-43}$$

is called Sommerfeld's fine-structure constant and plays an important role in the detailed spectral theory and quite generally in quantum physics. For instance, the velocity of the hydrogen electron on the first Bohr orbit of the H atom is 1/137 of the velocity of light. Our fourth dimensionless quantity [including (5-41)] contains the gravitational constant f and can be written

$$\frac{e^2}{mMf} = \gamma = 2.28 \times 10^{39} \tag{5-44}$$

It is the ratio of the electrostatic to the gravitational attraction between electron and proton, and thus also has a clear physical significance.

Using the four relations (5-41) to (5-44) we can, formally at least, express three of the six constants (5-38) by the three others, and thus go over from the cgs system with its arbitrarily chosen units, centimeter, gram, and second, to a more natural system of physical units. *We think that only h, c, and l_0 are suited as such natural units, because h is the characteristic constant of all quantum physics and c the characteristic constant of the relativity theory and any other field theory; the special importance of the fundamental length l_0 has been emphasized above.*

We therefore choose as reasonable new fundamental quantities the action h, the velocity c, and the length l_0. By using Eqs. (5-40) to (5-44) we express the constants e, m, M, and f by h, c, and l_0, and find

$$e = \sqrt{\frac{hc\alpha}{2\pi}} = 3.41 \times 10^{-2}\sqrt{hc} \tag{5-45}$$

$$M = \frac{h}{l_0c} \tag{5-46}$$

$$m = \frac{h}{\beta l_0 c} = 5.45 \times 10^{-4}\frac{h}{l_0c} \tag{5-47}$$

$$f = \frac{e^2}{\gamma m M} = 9.37 \times 10^{-40}\frac{l_0^2c^3}{h} \tag{5-48}$$

We now can express the old units, centimeter, gram, and second, by the new units h, c, l_0, and get

$$\left.\begin{aligned} 1\ \text{cm} &= 7.58 \times 10^{12}\, l_0 \\ 1\ \text{gm} &= 5.98 \times 10^{23}\,\frac{h}{l_0c} \\ 1\ \text{sec} &= 2.27 \times 10^{23}\,\frac{l_0}{c} \end{aligned}\right\} \tag{5-49}$$

or inversely

$$\left.\begin{aligned} &\text{The new unit of length } l_0 = 1.32 \times 10^{-13}\ \text{cm} \\ &\text{The new unit of mass } \frac{h}{l_0c} = 1.67 \times 10^{-24}\ \text{gm } (= M) \\ &\text{The new unit of time } \frac{l_0}{c} = 4.40 \times 10^{-24}\ \text{sec} \end{aligned}\right\} \tag{5-50}$$

In a formalistic way, we thus have reduced three of the six fundamental constants (5-38) to the three apparently most important ones and thus have found the relations between the old and the new physical units. It is an advantage of our choice of l_0 (5-40) that the new mass unit in the h, c, l_0 system automatically is the mass of the nucleon, or, more specifically, that of the proton.

The decisive problem, still to be solved, is how to derive, in a physically reasonable way, the dimensionless constants (5-42) *to* (5-44), *which must have a significance in physics and, consequently, should follow automatically from a rigorous and complete atomic theory* (*including a theory of the elementary particles*). We expect that it is not merely accidental that we get, from the six constants (5-38), two values of a length, (5-39) and (5-40), which lie so close together. We can express the same fact by saying that M/m is approximately $2hc/e^2$.

In spite of a number of spirited suggestions of Eddington and others, a satisfactory theory of the physical significance of the dimensionless constants is still missing. Moreover, there exists not even an agreement on the most promising approach. Particular difficulties are expected from the constant (5-44), because it seems extremely improbable that a constant of the order 10^{39} follows from normal atomic or general physical relations. This has led to the attempt to regard this quantity as evidence for an internal relation between atomic and astronomical data. The pertinent considerations of Eddington, Dirac, Ertel, Jordan, and others will be of great interest to any unbiased reader and will sharpen his view on the possibility and physical meaning of such relations. The wide divergence between the viewpoints of the different authors, however, may be taken as evidence that up to the present time all these combinations are no more than interesting, but very vague, hypotheses. These considerations show clearly, on the other hand, that, *besides the extension of quantum mechanics and the theory of elementary particles, the explanation of the dimensionless constants of physics, Eqs.* (5-41) *to* (5-44), *presents the third fundamental problem of modern physics. There seems to be little doubt, however, that these three problems are closely interrelated.*

Literature

Bethe, H. A.: Elementary Nuclear Theory, John Wiley & Sons, Inc., New York, 1947.

Bethe, H. A., R. F. Bacher, and M. S. Livingston: Nuclear physics, *Rev. Mod. Phys.*, **8**, 82, **9**, 69 (1936–1937).

Dänzer, H.: Einführung in die theoretische Kernphysik, Braun, Karlsruhe, 1948.

Eddington, A. S.: Relativity Theory of Protons and Electrons, The Macmillan Company, New York, 1936.

Feather, N.: An Introduction to Nuclear Physics, Cambridge University Press, London, 1936.

Feather, N.: Radioactive branching, *Rep. Progr. Phys.* **11**, 19 (1948).

Gamow, G.: Structure of Atomic Nuclei, Oxford University Press, New York, 1937.

Gentner, W., H. Maier-Leibnitz, and W. Bothe: Atlas typischer Nebelkammerbilder, Springer-Verlag, Berlin, 1940.

Goodman, C.: The Science and Engineering of Nuclear Power, Addison-Wesley Press, Cambridge, Mass., 1949.

Hahn, O.: Applied Radiochemistry, Cornell University Press, Ithaca, New York, 1936.

Heisenberg, W.: Die Physik der Atomkerne, 2d ed., Vieweg, Brunswick, 1949.

Heisenberg, W.: Vorträge über kosmische Strahlung, Springer-Verlag, Berlin, 1943.

von Hevesy, G., and F. A. Paneth: A Manual of Radioactivity, Oxford University Press, New York, 1938.

Hoag, J. B.: Electron and Nuclear Physics, D. Van Nostrand Company, Inc., New York, 1938.

Kamen, M. D.: Radioactive Tracers in Biology, Academic Press, New York, 1947.

Kopfermann, H.: Kernmomente, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1940.

March, A.: Natur und Naturerkenntnis, Springer-Verlag, Vienna, 1948.

Mattauch, J., and S. Flügge: Nuclear Physics Tables, Springer-Verlag, Berlin, 1942. (Translated in 1946.)

Millikan, R. A.: Electrons, Protons, Photons, Neutrons and Cosmic Rays, 2d ed., Chicago University Press, 1947.

Powell, C. F., and G. P. S. Occhialini: Nuclear Physics in Photographs, Oxford University Press, New York, 1947.

Rasetti, F.: Elements of Nuclear Physics, Prentice-Hall, Inc., New York, 1936.

Riezler, W.: Einführung in die Kernphysik, 3d ed., Bibliogr. Institut, Leipzig, 1944.

Rosenfeld, L.: Nuclear Forces, Interscience Publishers, New York, 1948.

Rutherford, E., J. Chadwick, and C. D. Ellis: Radiations from Radioactive Substances, Cambridge University Press, London, 1930.

Seaborg, G. T., and I. Perlman: Table of isotopes, *Rev. Mod. Phys.* **10,** 585 (1948).

Smyth, H. D.: Atomic Energy for Military Purposes (Smyth Report), Princeton University Press, Princeton, N.J., 1946.

Swann, W. F. G.: Cosmic rays, *Rep. Progr. Phys.* **10,** 1 (1946).

von Weizsäcker, C. F.: Die Atomkerne, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1937.

CHAPTER 6

MOLECULAR PHYSICS

6-1. The Goal of Molecular Physics and Its Relation to Chemistry

In Chap. 3 we acquired a fairly thorough understanding of the structure, properties, and behavior of atoms. In this chapter we shall discuss the structure of aggregates of atoms and their properties and processes, i.e., physics of molecules. The logical continuation of molecular physics is the physics of the liquid and solid state of matter, to be discussed in Chap. 7.

Molecular physics is the science of the structure and properties of molecules, so far as they are determined by *physical methods*. In this sense it is a logical continuation and extension of atomic physics proper. From what has been said it is obvious that molecular physics is related to chemistry. Chemistry attempts to determine the composition of a compound and the formula of its structure (e.g., CH_3Cl). From such chemical formulas, the chemist derives conclusions as to the behavior of the molecule, i.e., the possibility of its reaction with other molecules or atoms. The correctness of the chemical formulas is checked by chemical analysis, and finally methods are developed for synthesizing the molecules from their constituent elements. Chemical methods are also used to determine molecular properties and certain characteristic quantities, among them the heat of formation of the molecule, e.g., the energy required to form HCl from $\frac{1}{2}$ mole H_2 and $\frac{1}{2}$ mole Cl_2.

Molecular physics on the other hand attacks the molecule with physical methods. It determines the spatial arrangement of the atoms and the distances between them, its possibilities of rotation, and the moments of inertia of the molecule referred to different axes of rotation, its modes of vibration, and the dissociation energies required to split a diatomic molecule or to separate an atom or atom group from a polyatomic molecule. Last but not least it determines the arrangement of the electrons in the shells of the molecule, the possibilities of exciting electrons, and the effect of the excitation as well as ionization on the properties of the molecule. There is actually no known property of a molecule which is not tested and studied carefully by the molecular physicist.

Beyond that, however, it is an important task of molecular physics to determine how the atoms are bound in a particular molecule. The beginning student of chemistry often asks himself the question why there is a stable molecule of NH_3 but not of NH, a stable CO and CO_2 but no CO_4. It is an achievement of atomic physics with its quantum-mechanical atomic theory that we have now a possibility of answering this fundamental question by a theory of chemical binding. It is, therefore, not an overstatement when we say that only atomic and molecular physics has furnished a solid theoretical foundation for its older sister science, chemistry.

The foundations of molecular physics rest, as we hardly need to say, on atomic physics. The properties of atoms which were described in Chaps. 3 and 4 are prerequisite for understanding the structure of molecules. Therefore, the molecular physicist makes use of modern theoretical physics, especially quantum physics, to an increasing extent. He also utilizes, in his experimental studies, such specific physical methods as spectroscopy, X-ray and electron diffraction, the determination of dipole moments, or anisotropy measurements by means of the electrooptical Kerr effect. In general, chemical results are not used in molecular physics. The chemical behavior of a molecule should follow from the correct physical picture of the molecule, and it is used to verify the accuracy of this picture.

The results of just two decades of research in molecular physics are so striking, that modern inorganic and physical chemistry to an increasing extent not only has taken over many individual methods of molecular physics, but is actually doing molecular physics research just as it was done before by physicists. There is scarcely a border line any more between physics and chemistry, and this is especially welcome and necessary in the study of the polyatomic molecules of organic chemistry, for few physicists have that extent of chemical knowledge which is prerequisite for investigations in the field of organic molecular physics.

In our presentation we shall be concerned chiefly with the diatomic molecules, because their behavior is comparatively simple and easy to understand. Furthermore, fundamental research on diatomic molecules seems to have reached a certain state of completion. The polyatomic molecules will not be considered in detail, partly because research in this field is still in a fluid state, and partly because this field belongs more distinctly in the realm of chemistry. This treatment seems the more justified, as the behavior of polyatomic molecules, though much more complicated, does not differ too fundamentally from that of the diatomic molecules.

6-2. Methods of Molecular Research

In discussing the special methods of molecular research, we limit ourselves to purely physical methods. We thus shall disregard entirely the methods used in chemistry to determine the structure formulas.

The diameter of the molecule can be determined by methods discussed in the case of the atom on page 18, from the covolume (the constant b in the van der Waals equation of state), from the density in the liquid or solid state, as well as, for gases, from the measurement of viscosity. The statements made about the problem of defining the radius or diameter of the atom, page 18, hold also for the molecule. Therefore, different methods produce different values, and the viscosity method especially yields a molecular diameter which is temperature-dependent. The diameter of most diatomic molecules is about 3 to 4 A. However, we should not think of the molecule as spherical in shape, but regard the diatomic molecules as ellipsoids, whereas the shape of the polyatomic molecules, depending on their particular structure, has to be regarded in a rough approximation as rod-like, tetrahedral, etc.

X-ray and electron diffraction methods have been developed into important methods for investigating the size, nuclear separation, and binding strengths especially of polyatomic molecules. In a manner which is not simple theoretically or easy to understand, the type and spatial arrangement of the diffracting centers can be determined by measuring the angular distribution of X-rays or electrons scattered by molecules. By a high refinement of the experimental methods as well as of the methods of computation, it has been possible to determine not only the nuclear distances and thus the atomic structure of the molecule, but also the electron density distribution and, from it, the type and strength of the binding forces between the different molecular groups or atoms of a complicated molecule (Fig. 201).

Bragg showed about 20 years ago that this method of "Fourier analysis of X-ray diffraction data" can be used to obtain a real picture of the arrangement of the diffracting centers, e.g., of the atoms of a complicated molecule. According to Abbe's theory, each optical image can be interpreted as a diffraction phenomenon. By comparing this theory with the X-ray Fourier analysis, Bragg came to the conclusion that the pictorial representation of a molecule or crystal (Figs. 163 and 201), obtained from the X-ray diffraction data, corresponds to what *one would actually see with a hypothetical X-ray microscope.* Nevertheless, it is misleading when pictures such as Fig. 163 are occasionally referred to in popular publications as a "direct photograph of an individual molecule." An X-ray diffraction pattern is a result of the fact that monochromatic X-rays, due to the interference of the individual rays, are scattered by the lattice points (atoms) only into definite directions in space and with characteristic intensities. Since there is a definite relation between the diffraction pattern and the arrangement of the scattering atoms in a molecule or elementary crystal cell (page 425), this arrangement can be determined by working backwards from the angles and intensities of the traces on the X-ray diffraction pattern. This can be done photographically in a manner which is not quite easy to understand. The so-called "photograph" of the molecule or crystal is obtained by printing a number of diffraction patterns of macroscopic diffraction gratings with different grating constants, different orientations, and different exposures on the same photographic film. The grating constants, orientations, and relative exposure times can be calculated from the X-ray diffraction data. Figure 163 shows the result of this rather indirect and complicated, but very useful, method for the hexamethylbenzene molecule, $C_6(CH_3)_6$. The hexagonal

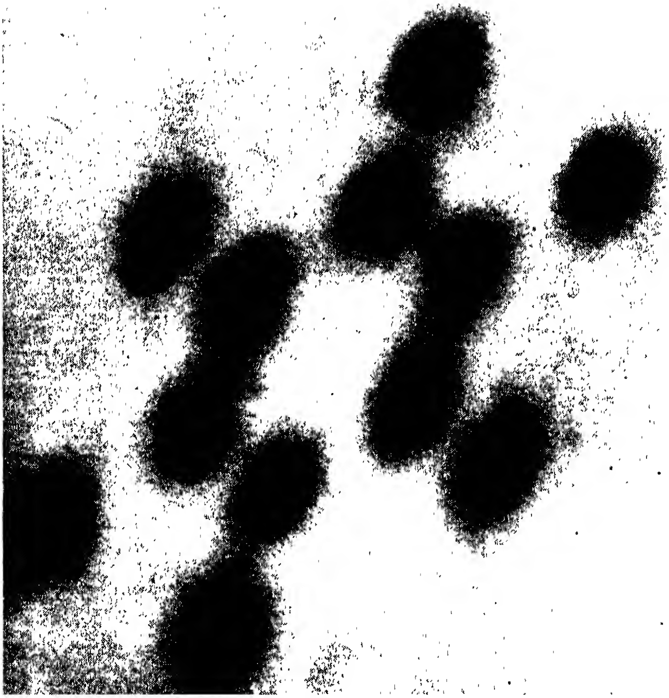


FIG. 163. Photographic construction of a picture of the hexamethylbenzene molecule, $C_6(CH_3)_6$, based on X-ray data by Brockway and Robertson. (*Courtesy of M. L. Huggins.*)

ring of the benzene molecule as well as the carbon atoms attached to its corners are clearly recognizable, whereas the H atoms, due to their small scattering power which is proportional to the square of the atomic number Z, do not show up.

Further information about the structure of molecules can be obtained from measuring their electric dipole moments, if they exist. These dipole moments, which can be computed from the measurement of the dielectric constants of gases or liquids, are of the order of magnitude of 10^{-18} esu. This is the order of magnitude to be expected if the displacement of the centers of the positive and negative charge, which are of the order of magnitude of the elementary charge (4.8×10^{-10} esu), is 0.1 A (10^{-9} cm). Dipole moments alone are not sufficient to draw conclusions about the structure of the molecule under consideration. However, they can be used to distinguish between different possible atomic arrangements obtained from other experiments. Thus we know immediately, for example, from the dipole moment of N_2O, that the symmetric shape of the molecule with the oxygen atom in the center is not possible, since in this case the dipole moment would be zero because of the symmetrical charge distribution. A more detailed discussion actually shows that the N_2O molecule has the unsymmetric form $N{\equiv}N{=}O$. In the physics of polyatomic molecules, the so-called physical stereochemistry, it is, in many cases, possible to coordinate electric moments with certain definite bonds in the molecule. The expected resulting electric moment for different possible molecular models can then be estimated from the contributions of the different molecular groups, if the mutual disturbance of different bonds affecting the same atom is taken into account. This resulting electric moment, computed for the different possible atomic arrangements, by comparison with the actually measured dipole moment, often allows one to determine the correct structure.

Measurements of the optical anisotropy of molecules by means of the Kerr effect can be a valuable supplement to dipole measurements. The electrooptical Kerr effect is that effect by which, in an electric field, certain molecular fluids become doubly refractive, so that they propagate light polarized in different directions with different velocities. The Kerr effect, as may be stated without going into the theory, depends on the difference in polarizability (production of a dipole moment by charge displacement in an electric field) along different axes of the molecule, and on the interaction, as a result of these different polarizabilities for different molecular axes, with the electric field vector of the two light waves polarized perpendicular to each other. Kerr effect measurements give us information about the polarizability of the molecules along different axes and thus quite generally about molecular structure.

All methods described here, if taken together, can give a good picture of the structure of a molecule and, in combination with the optical methods to be mentioned later, are especially valuable for explaining the structure of very complicated polyatomic molecules.

However, the most important evidence concerning the structure and the behavior of the simpler molecules has been obtained by spectroscopic methods, infrared spectroscopy, Raman spectroscopy, band spectroscopy, and recently, microwave spectroscopy. Spectroscopic research yielded, without using the above mentioned supplementary methods, a basic understanding even of the finest details of the behavior of the diatomic molecules, e.g., the interaction between electronic motion, vibration, and rotation. It has made possible the determination of all interesting molecular data with a precision which cannot be obtained with other methods. In the case of polyatomic molecules, spectroscopic research, though still in progress, has already given us information about the possibilities of vibration and dissociation and the influence of the excitation of electrons on them. Spectroscopic methods also permit us to determine, in many cases, the type and stability of different bonds, the geometrical arrangement of excitable electrons, and other details.

In all these studies, the investigation of the mutually supplementary absorption and emission spectra is rounded out in a fortunate way by the Raman effect, which permits us to determine the optically inactive vibration and rotation frequencies which cannot be determined from the other spectra. We add a few words about the Raman effect. *If monochromatic light of frequency* ν_0 *is scattered by molecules, there appear in the spectrum of the laterally scattered light some faint so-called Raman lines of frequency* $\nu_0 \pm \nu$*, where* ν *is equal to some vibration or rotation frequency of the scattering molecule, which thus can be determined by measuring the wavelength difference between the primary line and the Raman line.* This effect, which was predicted in 1923 by Smekal on the basis of the quantum theory and found by Raman and simultaneously by Landsberg and Mendelstam in 1928, can be easily explained if we consider only the energy balance. The photon, in being scattered by the molecule, gives up part of its energy in exciting higher rotational or vibrational states of the molecule, and the rest is scattered as a Raman line of lower frequency. Conversely, the shortwave, so-called anti-Stokes Raman line is due to a transfer of rotational or vibrational energy from the scattering molecule to the photon. To explain this process of nonclassical scattering of photons in a really satisfactory way which is in quantitative agreement with the experimental results and also with reference to the intensity of the Raman lines, we have to apply the quantum-mechanical theory of dispersion, which cannot be discussed here. It is sufficient to remark that *only such changes in energy states occur in the scattering process, i.e., only such vibrations are "Raman-active," whose excitation causes a change of the polarizability of the molecule. This is in contrast to the emission and absorption spectra in which only such energy changes occur by which the electric moment of the system is changed.*

Investigation of the longwave rotation spectra, of the rotational structure of rotation-vibration and band spectra, as well as the rotation Raman effect permits the exact determination of the moments of inertia of a molecule about its different axes of rotation. In the case of diatomic molecules with known masses of the constituent atoms, the internuclear distances can be computed from the moments of inertia for each electron and vibration state of the molecule. The investigation of the infrared vibration spectra and the vibrational structure of the band spectra, as well as the vibration Raman effect, enables us to compute in the case of the diatomic molecules the constants of the quasi-elastic binding forces, for each electron configuration, from the fundamental vibration frequencies. In the case of polyatomic molecules with many possibilities of vibration, however, the *modes of vibration* must be known in order to determine the constants of the binding forces from the different measured vibration quanta. Investigation of the vibrational structure of band spectra and especially the study of continuous molecular spectra furnishes the decisive data of its "breaking strength," the dissociation energy of the normal molecule as well as of the molecule in its different electron excitation states. These data, as well as the details of the dissociation processes under absorption of radiation, are of great importance for photochemistry. Important conclusions can also be drawn from continuous spectra about the process of formation of molecules from their constituent atoms. For all these studies, the spectroscopist has available not only the *stable* molecules known to the chemist but also normally unstable molecules. It has been especially valuable for the explanation of molecular structure, as for the kinetics of reactions, that radicals and unstable intermediate products such as OH, NH, C_2, diatomic metallic hydrides, etc., can be produced in sufficient concentrations in electric discharges, so that their spectra can be studied with the same accuracy as those of stable molecules.

The new *microwave spectroscopy* (see page 72) is being applied to molecular studies with surprising success. With its help, the measurement of very small energy differences can be carried out with high precision. According to page 388, for example, the quantized changes of the rotation energy are inversely proportional to the moment of inertia of the rotating molecule. The microwave methods give a means of determining the moments of inertia and, in simple cases, the internuclear distances and the atomic arrangement of such molecules whose rotational structure cannot be resolved by normal spectroscopic methods. Stark and Zeeman splitting of molecular lines can be measured precisely with weak electrostatic and magnetic fields, respectively. This knowledge is important for molecular theory. Also, electric dipole moments of normal as well as excited atoms can be determined from Stark-effect measurements with microwaves. A further field of application concerns the energy splitting of rotation levels (page 394) which is caused by the interaction of the molecular field and that of the electrons in motion, as well as by the interaction of the molecular field and the quadrupole moment of the nucleus. The results of the investigation of interactions of this kind seem to give information also about changes of the electron arrangement due to chemical valence saturation (page 409), thus providing an experimental approach to this important field of molecular physics. The determination of exact values of nuclear quadrupole moments as well as of the nuclear spins from molecular microwave spectroscopy establishes a new relation between molecular and nuclear physics. The measurement of small energy differences such as those associated with the pendular motion of the N atom in the ammonia molecule NH_3 through the plane of the three H atoms (inversion spectrum of NH_3), on the other hand, is of great interest to molecular physics itself.

The same spectroscopic equipment which was mentioned on page 68 can be used for recording molecular spectra. Emission and absorption spectra can be used to supplement each other in a very convenient way. Fluorescence, i.e., the emission of light by molecules upon absorption of exciting radiation, usually of monochromatic light, often is a great help in spectroscopic investigations of molecules.

The recording of molecular absorption spectra is not fundamentally different from that of atomic spectra (page 67), so we need not discuss it further.

The conditions for intense excitation of molecular spectra, however, differ considerably from those for the excitation of atomic spectra. For atomic spectra strong excitation is favorable. To produce band spectra, the excitation must be "soft" enough (current densities sometimes of only a few milliamperes per square centimeter) to prevent dissociation, because otherwise the molecules which emit the band spectra will be split up into atoms or groups of atoms. The same conditions hold for thermal excitation. Thus only line spectra of *atoms* appear in the axis of the electric arc with its temperature of 6,000° K and more, while the cooler envelope of the arc serves as an intensive source for *band spectra.* By using an arc in air, the oxide and nitride spectra of almost all elements can be excited by introducing only a trace of the element into the arc. In a hydrogen atmosphere, similarly, a large number of hydride spectra can be obtained. The various forms of glow discharges with or without electrodes (electrodeless ring discharge) are well suited for exciting molecules of gases or vapors. The band spectra of neutral molecules are found in the positive column of the discharge, while in the negative glow the molecular ions such as CO^+, N_2^+, O_2^+, etc., are excited.

The excitation of band spectra by fluorescence, in which the molecule is irradiated with suitable shortwave (monochromatic) light, is of interest for many problems. The fluorescent radiation is observed perpendicularly to the direction of the exciting radiation. By varying the wavelength and thus the energy of the exciting radiation, under certain conditions one part of a molecular spectrum after the other can be excited. The absolute position of the corresponding terms in the term diagram thus can be established. Our knowledge of the molecules I_2, S_2, Hg_2, Cd_2, and Zn_2, for example, depends almost entirely on such fluorescence studies.

Figure 164 gives a survey of the various forms of band spectra of molecules. Figure 165 is an example of a Raman spectrum.

6-3. General Review about the Structure and Meaning of Molecular Spectra

If the emission and absorption spectra of an arbitrary molecule built up of two dissimilar atoms is investigated with the spectroscopic equipment described on page 68, three distinct groups of spectra are distinguished clearly, which differ with respect to their locations in different spectral regions and by their different degrees of complexity.

For diatomic molecules, a series of equidistant lines is found in the far infrared. Polyatomic molecules show in this region some relatively simple sequences of lines. For reasons to be discussed below, these spectra are called the *rotation spectra* of the molecules.

In the near infrared between 1 and 10 μ a number of line sequences are found which are obviously arranged in a regular manner. The sequences with the shortest wavelengths extend into the photographic infrared, e.g., Fig. 164*e*. These sequences of lines are called the *rotation vibration* spectrum.

Finally, there appears in the photographic infrared, in the visible

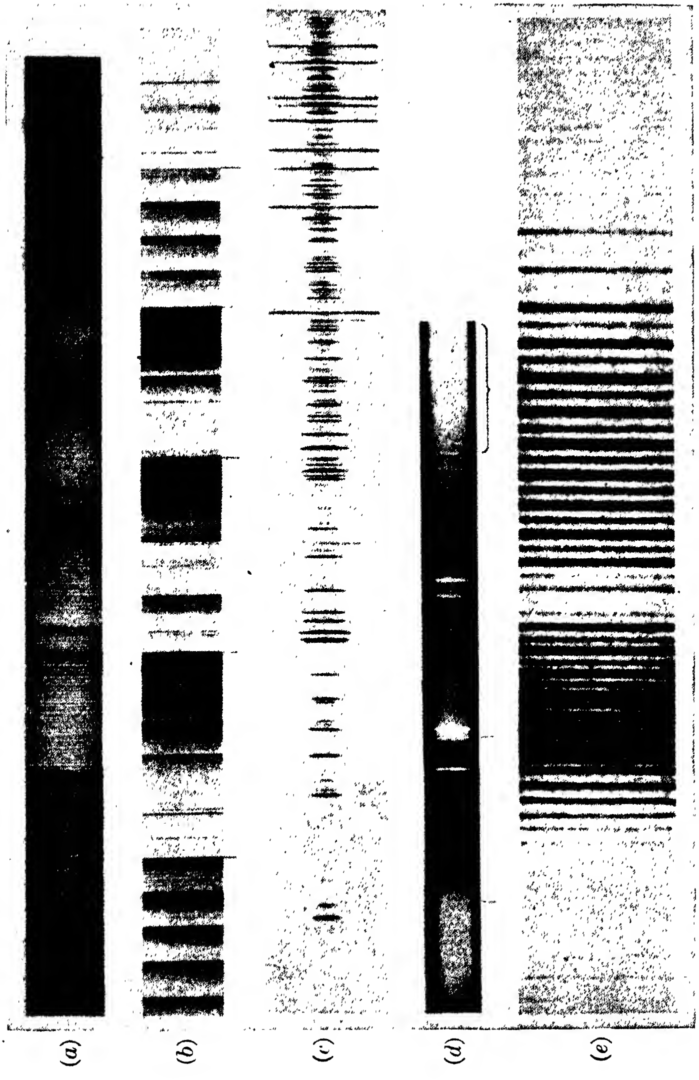


FIG. 164. Examples for molecular band spectra: (*a*) absorption band spectrum of the iodine molecule, photographed by Mecke with high dispersion (Rowland grating); (*b*) band spectrum of the PN molecule after Curry and Herzberg; (*c*) widely separated and therefore scarcely recognizable bands of the so-called many-line emission spectrum of the H_2 molecule (large Rowland grating, Fig. 28), photographed by the author; (*d*) emission spectra (diffuse bands and continua) of the Hg_2 van der Waals molecule (see page 385), photographed with low dispersion by Mrozowski; (*e*) absorption band of the C_2H_2 molecule, photographed with high dispersion (Rowland grating, Fig. 28) by Mecke.

and ultraviolet regions of the spectrum, a more or less large number of regularly arranged, but very complicated groups of lines (so-called *bands*). In addition to these bands, continuous spectra (see Fig. 164*d*) are also frequently found. These are called *band spectra* in a more restricted sense. The line sequences which appear here as in most molecular spectra, and which are shaded toward one side, are called bands, because they appear as structureless band-like phenomena in low dispersion spectroscopes. Each such band spectrum, lying in the shortwave spectral region, shows, in general, a threefold structure.

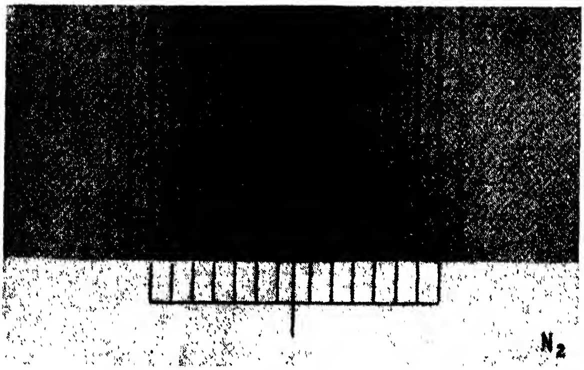


FIG. 165. Raman spectrum (rotational Raman effect, page 351) of the N_2 molecule. (*After Rasetti.*)

It consists of a number of clearly separated groups of bands, which, from their general appearance, apparently belong together, the so-called *band systems*. Each band system is built up of a greater or smaller number of *bands* which are occasionally arranged in sequences of bands, Fig. 164*a*. Finally, each band consists of a number of regularly arranged *band lines*.

This threefold structure of a band spectrum corresponds to a partition of the total energy of the molecule into three parts, the potential energy of the electron arrangement (electron energy), the energy of vibration of the atoms or of the nuclei of the molecule with respect to each other (vibration energy), and the purely kinetic energy of rotation of the molecule about a rotation axis perpendicular to the line joining the nuclei (or in the case of polyatomic molecules, about the three principal axes of inertia), *the rotation energy.* These three fractions of the total energy can change individually or together upon emission or absorption of radiation, and thus give rise to the different spectra.

The theory which we shall discuss later shows that the relatively small amounts of energy which correspond to the emission or absorption of the lines in the far infrared are due solely to changes of the rotation energy of the molecule. The spectrum associated with these changes of the rotation energy is called the *rotation spectrum*. The spectrum lying in the near infrared corresponds to changes of the vibration *and* rotation of the molecule; the superposition of these two types of energy changes produces the numerous lines and the complexity of the *rotation vibration spectra*. The band spectra lying in the shortwave region, i.e., in the visible and ultraviolet, correspond to changes in the configuration of the electrons, the vibration *and* the rotation of the molecule. These simultaneous, threefold energy changes account for the often very complicated structure of the *electron spectra* (band spectra in the more restricted sense).

The largest change in the total energy of the molecule occurs as a result of a change in the configuration of the electrons, i.e., by an electron transition. Thus the position of the band system in the spectrum is determined by the magnitude of the electron transition. All bands of a band system, therefore, belong to the same electron transition. The change in molecular energy due to a change in electron states is of an order of magnitude greater than that due to the quantized changes of the vibration. The position of a band in the band system thus is determined by the vibrational transition, i.e., the difference of the vibration energy in the initial and final state. The different lines of a band are the result of quantized changes of the rotation of the molecule. The change in the electron configuration, i.e., the electron transition, as well as the change in the vibration state are the same for all lines of the same band. The change of the total molecular energy as a result of a rotational transition is again of an order of magnitude smaller than the change in energy corresponding to a vibrational transition.

We can treat this same problem of the structure of a band spectrum from a somewhat different standpoint by considering the simplest model of a diatomic molecule. This model consists of two positively charged nuclei separated by a distance r, and an electron shell which in some way, yet to be determined, must hold the two nuclei together in spite of the electric repulsion between them. In the old pictorial model of a molecule, the outer electrons revolve around *both* nuclei. This idea of a common electron shell is in good agreement with the kinetic theory, according to which a diatomic molecule acts in a collision as if it were an ellipsoid of rotational symmetry with a relatively small eccentricity (Fig. 166). If one imagines the internuclear distance in a diatomic molecule (Fig. 166) to approach zero with the two nuclei fusing into one, one obtains an atom. From this picture it follows *that the electron shells, and with them the term diagram and spectrum of a diatomic molecule,* *must have a definite relation to the electron shells, the term diagram, and the spectrum of the atom which is formed by fusion of the two atoms of the molecule.* The H_2 molecule thus must be similar to the He atom.[1]

Now if we assume that in our model of the molecule, Fig. 166, changes of the energy state occur, as they do in the atom, only by electron transitions, then the spectrum would consist of individual lines corresponding to the different electron transitions, the so-called *fundamental lines,* which would appear in the wavelength regions of the actually observed band systems. Now if we include the changes of vibration of the nuclei within the molecule as possible changes of energy states, then we superimpose on each electron transition a large number of possible changes of the vibration energy. Each fundamental line thus would split up into a system of so-called *zero lines* which would appear at the same places in the spectrum as the actual bands in the band system. As a last approximation to the actual state of affairs we must include the quantized changes of the rotation of the molecule about the rotation axis, shown as a dotted line in Fig. 166, as possible changes of the total energy. Then each zero line splits up into a whole band, i.e., a series of *band lines,* because to each quantized change of the electron and vibration energy the multiplicity of the possible changes of the rotation energy has to be added. In this way we obtain the actually observed threefold structure of a band spectrum, though this is not always immediately apparent in a band spectrum (compare the spectra, Fig. 164).

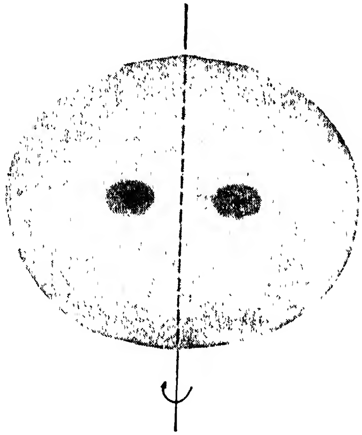


FIG. 166. Schematic representation of nuclei and electron cloud of a diatomic molecule (with rotation axis).

We shall discuss in the following, first, the motion of electrons in the molecule, i.e., the systematics of their electron states, then the molecular vibration and the phenomena related to it, and third, the rotation of the molecule and its influence on the spectra. This then leads to the discussion of the interaction of electronic motion, vibration, and rotation, from which the complete band spectra result.

---

[1] That this picture of the He atom, originating from the fusion of the two H atoms of the H_2 molecules, pertains only to its electron shells and nuclear charge but not to the nuclear mass, is unimportant since the two missing neutrons have almost no influence on the electron shells and their energy states and spectra.

6-4. The Systematics of Electron Terms of Diatomic Molecules

In this discussion of the electron states of diatomic molecules we shall start from the corresponding treatment of atomic states in Chap. 3. Reference is made to the corresponding sections of that chapter for the foundation of the quantum numbers and all details mentioned therein.

First we investigate the behavior of electrons in the so-called two-center molecular model, i.e., in the electric field of two positive nuclei rigidly held at some definite distance r (not necessarily identical with the equilibrium distance r_0 of the nuclei). Then we must determine what influence the change in internuclear distance, resulting from vibration and rotation, has on the electron arrangement. If we imagine the internuclear distance to be reduced until the nuclei coincide, we have a conversion of the electron shells and thus of the electron states of the molecule to those of an atom, whose nuclear charge is equal to the sum of the charges of the nuclei of the molecule. On the other hand, if the internuclear distance is increased to infinity, the electron levels of the molecule will go over continuously into those of the two atoms which actually are produced by dissociation of the molecule. Between these two limiting cases, in which we have atoms produced, in one case, by fusing the nuclei, and in the other by completely separating them, we can introduce a case which is already known to us. The electron configuration of the molecule is influenced by the axial electric field, whose direction is that of the molecular axis joining the two nuclei. The effect of this field on the configuration of the electrons and thus on the electronic states must be similar to that of an external electric field on the electrons of the atom produced by fusing the two nuclei. Its electron terms in the Stark effect must, consequently, be very similar to those of the molecule. Thus, we have the following similarity sequence of electron terms: combined-nuclei atom → atomic Stark effect → molecule → separated atoms. In this sequence, the energetic order and arrangement of the electron levels of the strongly bound normal molecules with a relatively small nuclear separation must be similar to that of the Stark effect of the atom formed by fusing the nuclei. On the other hand, weakly bound molecules with a large internuclear distance (molecules that are nearly dissociated, or van der Waals molecules, page 384) can be described in a better approximation by comparing their energy levels with those of the completely separated atoms. A rigorous development of this method of approach actually allows an accurate prediction of the electron term diagram of the molecule, as has been shown especially by Hund.

The behavior of molecular electrons, just as that of atomic electrons, can be quantitatively described by wave-mechanical eigenfunctions or, in a simpler but less accurate way, by four quantum numbers. As in the case of atomic electrons, the principal quantum number n represents the number of the electron shell to which the molecular electron under consideration belongs. The orbital quantum number l indicates the "form of the electron" (see Fig. 116) or, in the Bohr theory, the eccentricity of the orbit (see page 130). This is of less importance in molecular physics than in the atomic case. The orientation of the orbital momentum $\mathbf{l}$ which we have studied in the case of the Zeeman and Stark effects in atomic spectra, page 155, is caused, in the case of the molecule, by the axial field of the two positively charged nuclei. Thus the electron orbital momentum $\mathbf{l}$ precesses about the axis joining the nuclei with a quantized component in the direction of this axis. The quantum number λ corresponding to this orbital momentum component $\boldsymbol{\lambda}$ is given by the relation

$$|\boldsymbol{\lambda}| = \lambda \frac{h}{2\pi} \tag{6-1}$$

λ corresponds to the orientation quantum number m of the atomic electrons. Therefore, λ can have the $2l + 1$ different values

$$\lambda = l, l-1, l-2, \ldots, 0, -1, \ldots, -l \tag{6-2}$$

in which, in general, positive and negative values have the same energy. These states thus are degenerate, if the degeneracy is not removed by a perturbation, such as by the rotation of the whole molecule. The fourth electron quantum number, s, designates the spin direction and, as in the case of the atomic electrons, it can only have the values $\pm\frac{1}{2}$. The field with respect to which the spin is oriented again is that along the molecular axis (line joining the nuclei of the molecule).

Just as for the atomic electrons, the following relations hold

$$l \leqq n - 1 \tag{6-3}$$

$$|\lambda| \leqq l \tag{6-4}$$

Instead of the quantum numbers $l = 0, 1, 2, 3, \ldots$, the symbols s, p, d, f are used to designate the molecular electrons, just as in the case of the atom. In addition, instead of the quantum number $\lambda = 0, 1, 2, 3, \ldots$, the symbols σ, π, δ, φ are used. Just as for atomic electrons, the principal quantum number n is written before the electron symbol. Thus a $3d\pi$ electron is one which has the quantum numbers $n = 3$, $l = 2$, and $\lambda = 1$. Here, too, small letters are used to characterize individual electrons.

For molecules with many external electrons the behavior of the whole electron shell of the molecule follows, in a manner similar to that for atomic electrons (page 138*ff.*), from the vectorial combination of the momenta corresponding to the individual electrons. Because of the strength of the axial field in the direction of the line joining the nuclei, the coupling in the electron shells of the molecule is similar to the Paschen-Back effect of an atom in a strong magnetic field (page 162). The interaction between the orbital momenta $\mathbf{l}_i$ of the individual electrons, which combine to form the resulting total angular orbital momentum **L**, is smaller in the molecule than the coupling of the individual electrons to the nuclear field. The $\mathbf{l}_i$ of the outer electrons (the inner, closed shell electrons compensate each other and thus can be neglected in the first approximation just as in the case of atoms) precess individually about the molecular axis with quantized integral components $\boldsymbol{\lambda}_i$. These $\boldsymbol{\lambda}_i$ combine, additively or subtractively according to their direction, to give the resultant quantized momentum about the molecular axis characteristic for the electron states of the molecule, $\boldsymbol{\Lambda}$,

$$|\boldsymbol{\Lambda}| = \Lambda \frac{h}{2\pi} \tag{6-5}$$

Instead of the Λ-values 0, 1, 2, 3, . . . , the symbols Σ, Π, Δ, and Φ are used to designate the corresponding term series of the molecule.

The spin momenta $\mathbf{s}_i$ of the electrons combine vectorially, as they do for atoms, to form the resultant spin **S** of the electron shell, which then precesses about the direction determined by the molecular axis, so that its components in the direction of $\boldsymbol{\Lambda}$, designated by $\boldsymbol{\Sigma}$, take on all integral values between $+S$ and $-S$. As a result of the magnetic coupling between $\boldsymbol{\Lambda}$ and the resultant spin **S**, similar to the result of the coupling between **L** and **S** in the atom, each term belonging to a definite value of Λ splits into a term multiplet of $2S + 1$ terms, which are distinguished by the quantum numbers $\Omega = \Lambda + \Sigma$ of the resultant angular momentum of the electron shells about the molecular axis. However, *in contrast to the corresponding total momentum* **J** *of the electron shells of the atom,* $\boldsymbol{\Omega}$ *is not* [*even if we disregard the small nuclear spin* (*page 182*)] *the total momentum of the molecule, because the rotation of the whole molecule contributes appreciably to the total angular momentum of the molecule.* Details about the vectorial combination of $\boldsymbol{\Omega}$ and molecular rotation follow on page 394.

Just as for the atom, the value of the multiplicity of the term $(2S + 1)$ is written as a superscript to the left of the Λ-symbol, and the value of the resultant angular momentum Ω as a subscript on the right. The four components of an electron state with $\Lambda = 2$ of a three-electron molecule with $S = \frac{3}{2}$ thus are designated by ${}^4\Delta_{\frac{1}{2}}$, ${}^4\Delta_{\frac{3}{2}}$, ${}^4\Delta_{\frac{5}{2}}$, and ${}^4\Delta_{\frac{7}{2}}$. Figure 167 shows the vectorial combination for this case. To completely characterize an electronic term of a molecule, the symbols of the individual electrons are written before the symbol of the resulting molecular state. The ground state of the H_2 molecule with two $1s\sigma$ electrons thus is written $(1s\sigma)^2\ {}^1\Sigma_0$. The multiplicity law (page 134) holds for the molecule as it does for the atom. It states that a molecule with an even electron number has an odd multiplicity and vice versa. This law can occasionally be used to distinguish between spectra of molecules and their ions (e.g., N_2 and N_2^+) which otherwise may be difficult to separate, since the ions have one electron less than their neutral molecules and thus their spectra must have different multiplicity.

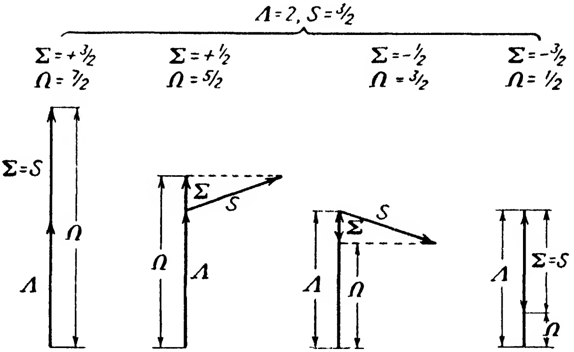


FIG. 167. The different possibilities of orientation of the angular momenta belonging to the molecular quantum numbers S and Λ in the case of a ${}^4\Delta$ term.

The selection rules for electron transitions in molecules correspond to a certain extent, but not entirely, to those of the atom; the difference is caused by the influence of the axial field. The principal quantum number n can change by an arbitrary integer, whereas the selection rule for Λ is

$$\Delta\Lambda = 0 \text{ or } \pm 1 \tag{6-6}$$

The selection rule for the spin is the same as for the atom: spin changes in the case of optical transitions, i.e., intercombinations between term systems of different multiplicity, are strongly forbidden for light molecules such as H_2 (just as for light atoms, page 140 and page 147). But they do occur for heavy molecules just as for heavy atoms, though with relatively small intensities. The so-called atmospheric oxygen bands, a ${}^3\Sigma \rightarrow {}^1\Sigma$ transition of O_2, and the Cameron bands of CO, a $^3\Pi \rightarrow {}^1\Sigma$ transition, are observed examples of such intercombination bands.

The quantum numbers Λ and Ω of a molecular term as well as its multiplicity can be determined empirically from the band spectra. In doubtful cases the Zeeman effect (page 159) plays a decisive role as a criterion, just as it does for atomic spectra.

Just as for atoms, Rydberg series of band systems are found for diatomic molecules with not too many outer electrons. From this it is concluded that generally a large amount of excitation energy is used, not to excite several electrons, but preferably for the higher excitation of *one* electron. Particularly, Price found in the vacuum ultraviolet Rydberg band series for N_2, O_2, CO, and NO (as well as for numerous polyatomic molecules). From these the spectroscopic value of the ionization energy of the molecules was determined by extrapolating to the series limit.

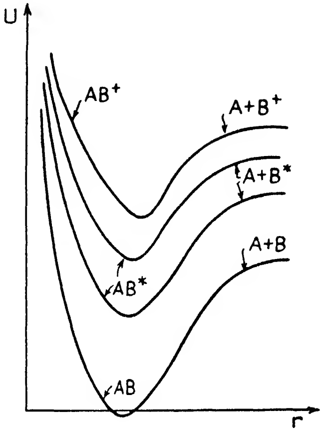


Fig. 168. Potential curves of a normal molecule, an excited molecule (two different excitation states), and the molecular ion for the case of very small influence by the excited electron on the molecular bond. On the right side the dissociation products for the different molecular states are indicated.

Such Rydberg series evidently occur only in certain molecular spectra. The reason for this becomes clear when we discuss the dependence of the potential energy of the electronic states of a molecule on the internuclear distance, and the influence of an electron excitation on molecular binding. Several potential curves of molecules are plotted in Fig. 168, which show that for a given electron arrangement the potential energy of the molecule is a function of the nuclear separation. We shall return to the details of this concept of potential curves on page 366. These potential curves are plotted in Fig. 168 for the ground state of the molecule, AB, for two excited states, AB^*, and for the ground state of the molecular ion, AB^+. *Only curves with a pronounced potential minimum correspond to a stable molecular state.* Thus for the case shown in Fig. 168, the molecules and molecular ions (and also the intermediate states of the excited molecule) are stable. In this case the excitation and finally the separation of an electron has *no* essential influence on the binding of the atoms, and Rydberg series can occur.

There are two other cases which occur and which are important for the understanding of molecular structure. They are those cases for which either the molecule or the molecular ion are unstable. Instability, i.e., a potential curve such as that of the ground state in Fig. 170, means that for any arbitrary separation between the atoms the repulsive forces predominate. Consequently, potential curves of this type (of unstable molecules) are called *repulsion curves.* It is possible that a stable molecule becomes unstable by losing an electron, so that the potential curve of the molecular ion has no minimum. Figure 169 shows

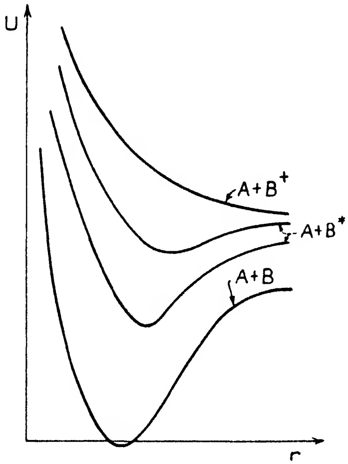


FIG. 169. Potential curves of a normal molecule, excited molecule (two different excited states), and ionized molecule for the case of strongly binding influence of the excitable electron.

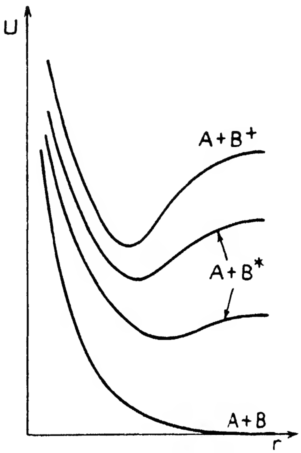


FIG. 170. Potential curves of an unstable normal molecule, excited molecule (two different states), and ionized molecule for the comparatively rare case of a loosening effect of the excitable electron on the molecular bond (for example, He_2).

the potential curves for this case. We conclude from it that the excitable electron was essentially responsible for the binding. For the third case there are, for example, the noble gases, which do not form normal molecules in the ground state, but which do have stable ions. This is shown in Fig. 170. When an electron approaches a molecular ion (the molecular core), the ion loses its stability so that a stable molecule does not exist in the ground state: we speak of this as a loosening electron in contrast to the binding electron of Fig. 169. Since the excited molecular states represent, to some extent, intermediate steps between the normal molecule and the ion (Figs. 168 to 170), one can arrive at the value of molecular constants of unstable molecules or ions by studying the change of the spectroscopically determined molecular constants with increasing excitation (principal quantum number n).

6-5. Vibrations and Vibration Spectra of Diatomic Molecules

a. Vibration Levels and Potential Curves

We investigate now the vibration of the atoms or, since the light electrons do not play a very important role, of the nuclei with respect to each other, and the spectroscopic phenomena dependent upon molecular vibrations. As always, we start with the simplest model and gradually extend it. Thus we assume there are two point masses, the nuclei, which can vibrate only along a straight line joining them. Furthermore, the forces between the nuclei may be a function only of the distance r between them. For the case $r = r_0$ the repulsive and attractive forces may be in equilibrium and the potential $U(r)$ thus should have a minimum.

If we choose the origin of the coordinates to be in one of the nuclei (the one on the left) then we can treat the vibration of both nuclei about the center of mass as if the second nucleus alone would vibrate with respect to the origin located in the first nucleus. We introduce as a variable the relative internuclear distance, measured in units of the equilibrium distance r_0,

$$\rho = \frac{r - r_0}{r_0} \tag{6-7}$$

The reduced mass μ and the moment of inertia I of the molecule can be computed from the masses m_1 and m_2 of the two atoms and their equilibrium distance according to

$$I = \mu r_0^2 = \frac{m_1 m_2}{m_1 + m_2} r_0^2 \tag{6-8}$$

For the potential energy we use an expression corresponding to the harmonic force equation (4-72), page 219,

$$U(r) = E_{\text{el}} + b\rho^2 \tag{6-9}$$

Then we obtain from the general Schrödinger equation (4-42) the special Schrödinger equation (4-75), page 220, of the linear harmonic oscillator. In our notation this can be written as

$$\frac{d^2\psi}{d\rho^2} + \frac{8\pi^2 I}{h^2}(E_v - b\rho^2)\psi = 0 \tag{6-10}$$

Here E_v is the vibration energy of the molecule obtained by subtracting from the total energy E the energy of the electron configuration E_{el}. According to page 220 this equation has solutions only for the discrete energy eigenvalues

$$E_v = hc\omega(v + \tfrac{1}{2}) \qquad v = 0, 1, 2, \ldots \tag{6-11}$$

where

$$\omega = \frac{1}{c}\sqrt{\frac{b}{2\pi^2 I}} \tag{6-12}$$

is the wave number, measured in cm^{-1}, which corresponds to the natural frequency $\nu = c\omega$ of the molecule.

The energy level diagram of a vibrating diatomic molecule thus consists of equidistant steps of magnitude $hc\omega$. However, the lowest energy value is not zero but, according to (6-11), is $hc\omega/2$. v is called the vibration quantum number.

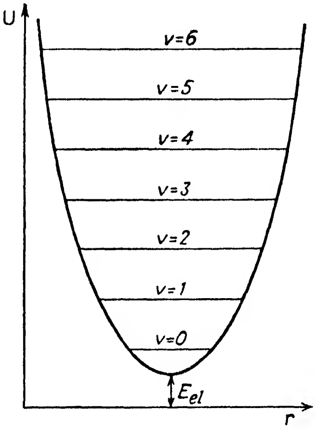


FIG. 171. Potential curve and vibration states for the harmonic oscillator.

To illustrate and further study the molecular vibration, we use the potential curve of Fig. 171. It shows the potential of the forces acting between the two nuclei as a function of the nuclear separation. The quantized vibrational energy states E_v are also indicated. The potential curve of the linear harmonic oscillator, according to (6-9), is simply a parabola. This potential curve can be interpreted easily by an approximate mechanical analogy. If one nucleus is imagined to be fixed at the zero point, then the other nucleus oscillates back and forth, just as a ball would roll up and down in the potential trough. The point of reversal of the vibration is the intersection of the discrete level of the vibration energy with the potential curve. At this point, all the energy of the molecule is potential energy.

There are two points in which the model of the harmonic oscillator does not agree with the actual molecule. First, the nuclear separation cannot be reduced to zero, because the electrostatic repulsion between the positive nuclei (and with it the potential energy U) rapidly becomes very large as the distance between the nuclei is reduced. Furthermore, the inner electron shells of the atoms in the molecule offer much resistance to any distortion. Conversely, with increasing internuclear distance, the potential curve $U(r)$ cannot rise to infinity, since the binding forces due to the electrons which hold the nuclei together must decrease with increasing separation until finally the atoms are torn apart, i.e., the molecule is dissociated. The real case thus corresponds somewhat to the curve in Fig. 172, i.e., we must go from a harmonic oscillator to the molecular model of an anharmonic oscillator. The potential curve

$U(r)$ corresponds closely to the actual behavior of the molecule if it is represented by the expression of Morse

$$U(r) = D[1 - e^{-a(r-r_0)}]^2 \tag{6-13}$$

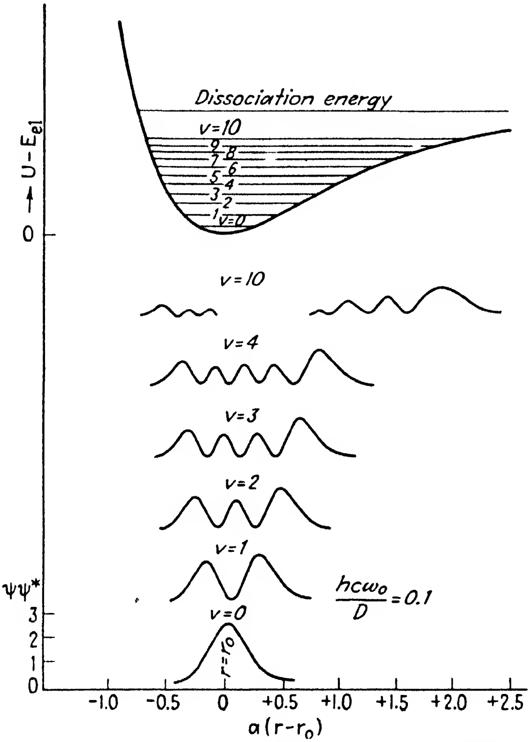


FIG. 172. Potential curve of an anharmonic oscillator corresponding to Eq. (6-13) with $\psi\psi^*$ functions for a number of vibration states. (*Computed by C. Gregory.*)

where

$$a = 2\pi\omega c\sqrt{\frac{\mu}{2D}} \qquad \mu = \frac{m_1 m_2}{m_1 + m_2} \tag{6-14}$$

The meaning of D will soon be discussed. If we introduce this $U(r)$ from (6-13) and (6-14) into the Schrödinger equation (6-10) the eigenvalues of the vibration energy E_v have, instead of (6-11), the form

$$E_v = hc\omega\left(v + \frac{1}{2}\right) - \frac{h^2c^2\omega^2}{4D}\left(v + \frac{1}{2}\right)^2 \tag{6-15}$$

In going over to the anharmonic oscillator we correct automatically for the second point in which the first model did not agree with experience. According to the spectra, the difference between consecutive vibration levels is not constant but actually decreases with increasing v values and converges toward a limit which is approached asymptotically by the right branch of the potential curve. The value of this energy limit turns out to be identical with the constant D of Eq. (6-13). Thus D is nothing else than the dissociation energy of the molecule. Figure 172 shows a potential curve with characteristic vibration levels computed by means of Eqs. (6-13) and (6-14). The norm $\psi\psi^*$ of the vibration eigenfunctions, as computed from the solutions of the Schrödinger equation, is plotted here for the vibration states $v = 0$, 1, 2, 3, 4, and 10. These energy levels and eigenfunctions have to be compared with those of the harmonic oscillator as shown in Fig. 114, page 221. We see that in Fig. 172 the maxima corresponding to large internuclear distances are larger than those for small separation. This means that, as a result of the anharmonicity of the molecular vibration, the nuclei are more likely to be found at a large internuclear distance than they would in a harmonic vibration.

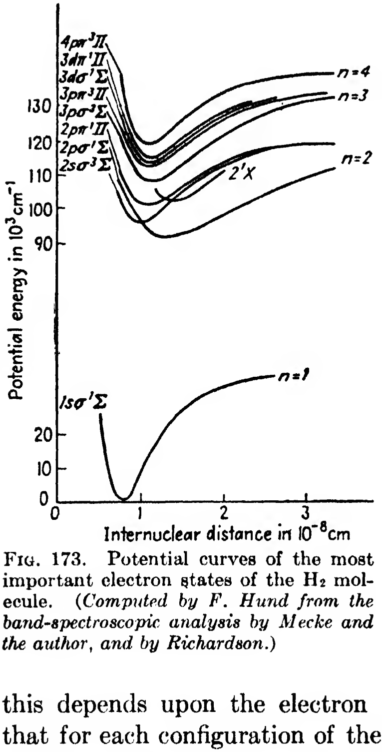


FIG. 173. Potential curves of the most important electron states of the H_2 molecule. (*Computed by F. Hund from the band-spectroscopic analysis by Mecke and the author, and by Richardson.*)

Since the shape and position of the potential curve is determined by the binding between the nuclei and this depends upon the electron configuration, it is easily understood that for each configuration of the electrons, i.e., for each of the excited electron states of the molecule, there is a characteristic potential curve. Now, according to page 364, the excitable electron generally participates in the binding, and its excitation loosens the bond. The equilibrium distance r_0 will, consequently, be greater and the dissociation energy smaller in an excited state than in the ground state. As an example, the actual potential curves of the H_2 molecule are shown in Fig. 173. These curves are computed from Eqs. (6-13) and (6-14) using the values of the dissociation energy D and the fundamental vibration quantum ω_0 taken from the spectrum of the H_2 molecule.

b. Change of Vibration States and Infrared Vibration Bands

We now inquire about the transitions between the vibration states which produce the emission or absorption spectra (bands). We have to distinguish between transitions within the same vibration level diagram, i.e., without change of the electron configuration, and vibrational transitions with simultaneous electron transitions. We treat the first case next. It accounts for the infrared rotation vibration bands if we take into consideration the molecular rotation which will be discussed later. *For the harmonic oscillator the selection rule for the vibration quantum number is* $\Delta v = \pm 1$. *This means that transitions under emission or absorption of radiation can occur only between neighboring vibration levels. For the anharmonic oscillator, which corresponds more closely to the actual molecular case, the selection rule is less stringent in so far as with decreasing probability (intensity) the transitions* $\Delta v = \pm 2, \pm 3, \ldots$, *are also possible.*

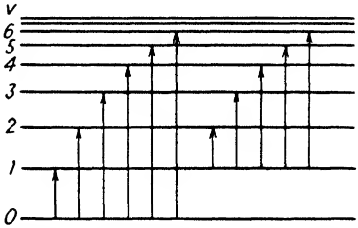


FIG. 174. Absorption transitions within a vibration-level diagram of the electron ground state of a molecule, leading to the vibration bands (though here without rotation!) of the molecule.

Now the energy difference of neighboring vibration levels of diatomic molecules is of the order of magnitude of 0.1 to 0.3 ev, i.e., it is considerably greater than the average thermal energy kT. Consequently, higher energy levels are normally *not* excited, so that the rotation vibration bands are observed almost exclusively in absorption. According to Fig. 174 only the bands starting from the ground state and, at higher temperatures, those from the first excited vibration state, are observed. Only very hot incandescent gases and vapors *emit* rotation vibration bands. The fundamental band, corresponding to the energy $\Delta E = 0.1$ to 0.3 ev, is found in the region from 4 to 12μ, whereas the higher energy transitions (harmonics) are found in the region up to the photographic infrared at about 1μ. In this region they can be investigated with high-dispersion spectrographs.

The emission and absorption of rotation vibration bands is restricted by a severe limitation. *An intense emission or absorption of radiation is possible only if the electric moment of the system changes as a result of the transition* (as in a classical dipole vibration). This is always the case for *electron* transitions because of the change in the distance between the electron and the nucleus. However, when identical atoms vibrate with respect to each other, e.g., in H_2, O_2, N_2, etc., the electric moment does not change. From this follows an important law (which is in agreement with our empirical knowledge) that *diatomic molecules consisting of identical atoms do not have observable rotation vibration bands.* These vibrational frequencies which do not appear in the optical infrared spectra are called *optically inactive* or *infrared-inactive* frequencies. However, since the polarizability changes in these cases with the vibration of the nuclei, these frequencies are *Raman-active,* see page 352. Thus they appear in the Raman spectrum and can be measured.

c. The Franck-Condon Principle as a Selection Rule for Simultaneous Electron and Vibration Transitions

We now discuss changes of the vibration state of molecules which are coupled with electron transitions. There is no stringent selection rule for such transitions between different vibration term systems, i.e., between potential curves of two different electron states of the molecule, Fig. 175. These are governed by the Franck-Condon principle which indicates the *preferred,* intensive transitions and thus the general appearance (intensity distribution) of all the electron band systems and the continuous spectra of a molecule. In discussing these transitions, Franck started with the reasonable assumption that, because of the small mass of the electron compared to that of the nuclei, the change of the electron configuration during an electron jump occurs so fast that the position and velocity of the nuclei does not change appreciably during the transition. This means that, *in the potential curves shown in Fig. 175, transitions occur without a change in the internuclear distance, i.e., vertically (with little scattering). Furthermore, because the velocity is preserved, transitions occur always between states for which the nuclei have the* same *velocity.* Since in the classical picture, see page 221, the nuclei remain longest at the reversal points of their vibration, there is the greatest probability for transitions to occur between two such reversal points. Transitions thus occur perpendicular to the energy level lines from and to the points where they intersect the potential curves. This is shown in Fig. 175. Condon verified this classical principle by wave mechanics. According to wave mechanics, page 216, the transition probability is given by the integral

$$\int \psi_i r \psi_f \, d\tau \tag{6-16}$$

in which ψ_i and ψ_f are the eigenfunctions of the initial and final states of the transition under consideration. From this it follows that the transition probability is large only for those combining states whose eigenfunctions have maxima for the same nuclear separation. If it is assumed, as a first approximation, that that part of the transition probability which is due to the electron transition is constant and independent of the nuclear separation, it follows from the wave-mechanical Franck-Condon principle that the preferred transitions occur between a maximum of the vibrational eigenfunction of one potential curve to a $\psi\psi^*$ maximum of the other potential curve. However, since, according to page 221 and Fig. 172, the maxima of the vibrational eigenfunctions almost coincide with the reversal points of the classical vibrations, the wave-mechanical computations lead to the same results as the classical picture. Only for the lower vibration levels is there a departure from the rule of perpendicular transitions between the reversal points, since we know from page 221 that here the maxima of $\psi\psi^*$ do not closely coincide with the classical reversal points but that they move toward the center. For the lowest level the maximum is almost exactly in the center. The fact *that the molecular spectra show an intensity distribution which is in excellent agreement with the form of the principle derived from wave mechanics, and not with the classical picture, is another indication of the correctness of wave mechanics and the necessity to use it in treating atomic and molecular physics.*

It is very important, especially for the theory of continuous molecular spectra, that the transitions do not occur exactly perpendicularly, but in a small region of internuclear distances whose width is determined by the finite width of the $\psi\psi^*$ maxima, Fig. 172. From Fig. 175 it can be seen immediately that it depends on the relative position of the potential curves, which of the vibrational transitions are preferred, i.e., which appear with large intensity. If the equilibrium distances for two electron states are the same, as in Fig. 175*a*, and if we let v' be the vibration quantum number of the upper and v'' of the lower electron state, then the bands

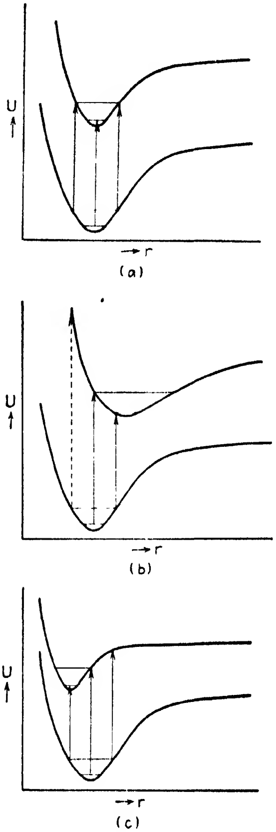


FIG. 175. Schematic representation of the transitions between ground state and excited electron state of a diatomic molecule, to be expected according to the Franck-Condon principle: (*a*) for identical internuclear distances in both electron configurations; (*b*) for loosening of the bond by electron excitation; (*c*) for tightening of the bond by electron excitation.

with $v', v'' = 0.0; 1.1; 2.2; \ldots$ appear with high intensity. If, on the other hand, r_0 is greater for the upper state than it is for the lower one (the usual case of weaker binding as a result of electron excitation, Fig. 175*b*), there are in the case of absorption *two* preferred transitions to the upper state for each v''. In the case of stronger binding in the upper state (Fig. 175*c*, where r_0 is smaller in the upper state than in the lower), there are two other preferred transitions.

This intensity distribution in a band system, which follows from the Franck-Condon principle, becomes clearer if a two-dimensional diagram is used to plot the transitions between two vibrational term systems. In the so-called edge diagram of a band system, the vibration quantum numbers v'' of the lower state are plotted as abscissas, and the v' of the upper state are plotted as ordinates. At the point corresponding to the quantum numbers v', v'', depending upon the purpose of the diagram, the wavelength, wave number, or intensity of the band corresponding to the transition $v'' \longleftrightarrow v'$ is written (Fig. 176). From the Franck-Condon principle it then follows that for the case shown in Fig. 175*a*, bands of high intensity lie on the diagonal of the diagram. In the two other cases they lie on a more or less open parabola. The bands of the vertical branch of this parabola appear in absorption, those of the horizontal branch in emission. The Franck-Condon principle thus explains why absorption and emission band spectra supplement each other so well. From an investigation of the intensity distribution in a band system by means of the Franck-Condon principle the relative positions of the potential curves of the two combining states can be determined. This, in turn, permits one to determine the influence of the excited electron on the binding strength (see page 364).

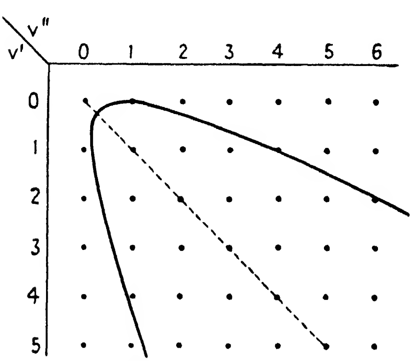


FIG. 176. Intensity distributions in a band system as expected according to the Franck-Condon principle. The dotted line shows the most intensive bands for the case of Fig. 175*a*; the solid parabola the most intensive bands for electron configurations corresponding to Fig. 175*b* or *c*.

d. The Structure of an Electron Band System. Band Edge Diagrams and Band Edge Formulas

Because the Franck-Condon principle is valid, we have at least in the two extreme cases (when the change in the nuclear separation is zero or very great) an entirely different appearance of a band system.

The two different cases are usually called "group spectra" and "sequence spectra," respectively.

Figure 177 shows the transitions which occur if the potential curves (Fig. 175*a*) have almost the same position. The spectrum then consists of the bands on the principal diagonal $\Delta v = 0$ and those on the two neighboring diagonals $\Delta v = \pm 1$. It can be seen from Fig. 177 that there is little difference in the length in reciprocal centimeters of the transition arrows for each of the three groups, and thus very little difference in the wavelengths of the corresponding bands. Consequently, the spectrum consists of groups of narrow bands lying close together. For this reason it is called *a group spectrum.* Typical group spectra are the band spectra of CN and C_2 which are emitted by every carbon arc. Conditions are different for the cases of Fig. 175*b* and 175*c*, i.e., for very different relative positions of the combining potential curves. Here, as can be seen from the parabola in the band edge diagram, Fig. 176, the *emission* spectrum consists mainly of a band sequence starting from the same upper state, Fig. 178, left side. In *absorption*, a similar band sequence is observed whose bands have the same lower state (Fig. 178, right side). Compared to the group spectra, the wavelength difference of neighboring bands is large; we have a *band sequence spectrum* such as that of the I_2 molecule (see Fig. 164*a*).

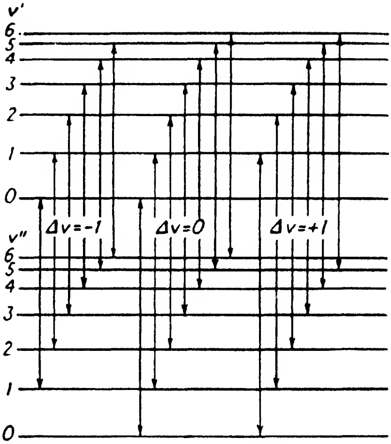


FIG. 177. The most intensive vibration bands in a "group spectrum" corresponding to the potential curves in Fig. 175*a*.

We now take up the question of how the vibrational structure, i.e., the vibration level differences in the upper and lower states, can be determined from an observed band system. We answer the question with the aid of the band edge diagram in which the wave numbers of the bands of the system (actually that of an edge or a line of every band) are noted. Table 13 (after Finkelnburg and Mecke) shows the first bands of a band system of H_2 in which the difference between two bands is noted in italics between the wave numbers of the two bands. According to the construction of the edge diagram (Fig. 176 and 178) all bands written in a horizontal row have the same initial state v' and those beneath have the next higher initial state $v' + 1$. The wave-number difference of the first two bands of each vertical column therefore must be constant and equal to the difference of the term values $v' = 0$ and $v' = 1$ in the upper state. Table 13 shows that these wave-number differences actually are constant up to a few thousandths of a per cent. Correspondingly, all bands in the same vertical column have the same end state: the differences between consecutive bands in the vertical columns are equal to the differences of the vibration terms in the lower state. The decrease of consecutive term differences (1,312 → 1,276 → 1,242 or 2,105 → 1,944 → 1,766) is a measure of the anharmonicity of the vibrations. From what has been said it seems clear that for arranging a band system in an edge diagram, the so-called vibration term analysis, the constant wave-number differences between bands are sought first and are then arranged in the right order in an edge diagram. The fact that the term differences, as a result of the anharmonicity of the vibration, must follow each other in a decreasing sequence, is a way of controlling the correctness of the band analysis.

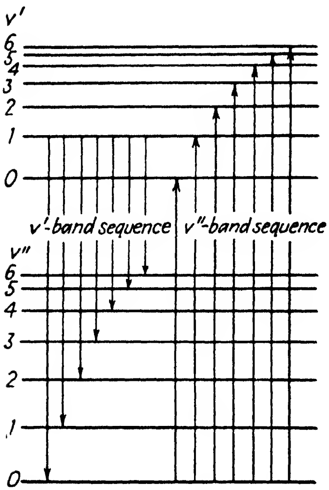


FIG. 178. Most intensive vibrational transitions (different in emission and absorption, see directions of the arrows!) in a "sequence spectrum," corresponding to the potential curves in Fig. 175*b* or *c*.

Table 13. Wave Numbers of a Band System of the H_2 *Molecule (incomplete) Arranged in an Edge Diagram (after Finkelnburg and Mecke). Wave-number Differences (in Italics) = Term Differences*

v' \ v''	0		1		2		3
0	21827.99	*1312.51*	20505.48	*1276.63*	19238.85	*1242.39*	17996.46
	2105.23		*2105.22*		*2105.26*		*2105.28*
1	23933.22	*1312.52*	22670.70	*1276.59*	21344.11	*1242.37*	20101.74
	1944.75		*1944.63*		*1944.74*		*1944.64*
2	25877.97	*1312.63*	24565.34	*1276.49*	23288.85	*1242.47*	22046.38
	1766.52		*1766.64*		*1766.53*		*1766.58*
3	27644.50	*1312.52*	26331.98	*1276.60*	25055.38	*1242.42*	23812.96

We now try to represent the empirically determined vibration structure by an analytical formula. We can then compare these empirical term formulas with the theoretical formulas for the vibration terms and thus determine important molecular constants. According to Eq.

(6-15) the wave number of a band which corresponds to the vibrational transition $v' \rightarrow v''$, is

$$\bar{\nu}(v',v'') = \frac{E'_v - E''_v}{hc} = \bar{\nu}_{\text{el}} + \omega'\left(v' + \frac{1}{2}\right) - \frac{hc\omega'^2}{4D'}\left(v' + \frac{1}{2}\right)^2 - \omega''\left(v'' + \frac{1}{2}\right) + \frac{hc\omega''^2}{4D''}\left(v'' + \frac{1}{2}\right)^2 \qquad (6\text{-}17)$$

For brevity we introduce

$$x = \frac{hc\omega}{4D} \qquad (6\text{-}18)$$

Since the actual anharmonicity is not represented with sufficient accuracy by the quadratic formula of our model, we supplement it by a cubic term in $(v + \frac{1}{2})$ including a new constant y and obtain

$$\bar{\nu}(v',v'') = \bar{\nu}_{\text{el}} + \omega'(v' + \tfrac{1}{2}) - \omega'x'(v' + \tfrac{1}{2})^2 + \omega'y'(v' + \tfrac{1}{2})^3 - \omega''(v'' + \tfrac{1}{2}) + \omega''x''(v'' + \tfrac{1}{2})^2 - \omega''y''(v'' + \tfrac{1}{2})^3 \qquad (6\text{-}19)$$

Here $\bar{\nu}_{\text{el}}$ is the wave number of the pure electron transition, i.e., the perpendicular distance between the potential minima of the combining states. However, the pure electron term never really appears; we can observe only the lowest vibration term, corresponding to $v = 0$ (the zero point vibration, page 220), which lies $hc\omega/2$ above the pure electron state. Computations in empirical band spectroscopy therefore are carried out with the term difference of the two lowest vibration levels; this term difference is denoted by $\bar{\nu}(0,0)$ instead of $\bar{\nu}_{\text{el}}$. Furthermore, band spectroscopists are accustomed to work with integral vibration numbers v instead of the quantum-mechanically correct half-integral values, because band spectra were analyzed empirically before the wave-mechanical theory was developed. The empirical band formula for representing the vibration structure of a molecule thus is written

$$\bar{\nu}(v',v'') = \bar{\nu}(0,0) + (\omega'_0 v' - \omega'_0 x' v'^2 + \omega'_0 y' v'^3) - (\omega''_0 v'' - \omega''_0 x'' v''^2 + \omega''_0 y'' v''^3) \qquad (6\text{-}20)$$

Its values can evidently be easily computed in terms of the above theoretical formula. As we have not yet discussed the rotation of a molecule, the band formula holds for the case where rotation = 0, i.e., for the so-called *zero lines* of the bands. However, the formula is used mostly to represent the band edges which alone can be measured with small dispersion spectrographs. Since the difference between the edge and zero line for the bands of one system is not constant, the band formula holds only approximately for the edges, but it does give (*if* the analysis is correct) as an *edge formula* a good representation of the actual vibration structure and it permits the determination of the vibration quanta and thus the binding force constants of the two combining molecular states.

6-6. Dissociation and Formation of Diatomic Molecules and Their Relation to the Molecular Continuous Spectra

a. Molecular Dissociation and the Determination of Dissociation Energies

In introducing the concept of the potential curve we mentioned several times the dissociation energy D and the possibility of a molecule dissociating as a result of excessive vibration. Now we want to treat in some detail this process and the problem of determining the dissociation energy which is one of the most important molecular constants. By dissociation energy we mean the energy necessary to completely separate the atoms of a diatomic molecule. However, physicists and chemists do not mean the same thing when they use the term dissociation energy. The chemist measures the dissociation energy of a mole of molecules and computes the value for the single molecule from that. In general, however, all molecules are not in the lowest vibration and rotation state, so that the chemical dissociation energy represents an average value. In order to agree with the dissociation energy obtained by the physicist from the band spectra of the molecule in the ground state (and for decomposition into *normal* atoms), the chemical measurements must be made in the neighborhood of absolute zero or be reduced to it. Besides the dissociation energy of a molecule in the ground state, the physicist is able to determine further dissociation energies which are not of interest to the chemist, i.e., the dissociation energies of the excited molecule. The dissociation here usually produces a normal and an excited atom. Therefore, in comparing spectroscopic and chemical dissociation energies, some caution must be used.

After this brief explanation we turn to the dissociation process itself. First, we must extend our present vibration level diagram. This extends not only to the dissociation limit, but beyond it toward higher energies (Fig. 179). Here there is a continuous energy region, corresponding to the states where the molecule is decomposed into its constituent atoms which may have varying amounts of kinetic energy. This continuous energy range (Fig. 179) beyond the dissociation energy of a molecule is analogous to that beyond the series limit (ionization energy) of an atom (page 106) which corresponds to the ionization of the atom into an ion and electron with varying amounts of kinetic energy.

One might be inclined to think that an optical dissociation should be possible by absorption of the dissociation energy (i.e., by overexciting the nuclear vibrations), while the electron remains in the ground state, i.e., without exciting the electron, according to the following equation

$$AB + hc\bar{\nu}_c \rightarrow A + B \tag{6-21}$$

However, this is not the case, because according to the selection rule mentioned on page 369, such great changes in the vibration quantum number, which correspond to the transition from the vibration ground state into the dissociation continuum, occur only with infinitely small probability. The same conclusion can be drawn from the Franck-Condon principle according to which the transition probabilities for such transitions are extremely small. As can be seen from Fig. 180, the maximum of the eigenfunction of the dissociated state under consideration does not lie above the maximum of the eigenfunction of the vibration ground state at all. *Thus an optical dissociation or photodissociation by a change* only *of the vibration state without a simultaneous electron excitation is not possible.*

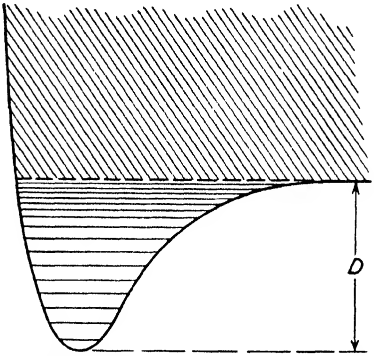


FIG. 179. Potential curve of a diatomic molecule with discrete and continuous energy states. D indicates the dissociation energy.

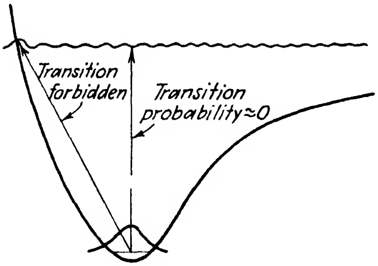


FIG. 180. Potential curve with vibration eigenfunctions for the vibrational ground state and a dissociated state, indicating the impossibility of optical dissociation of a molecule without simultaneous electron excitation.

The case is different for thermal dissociation. If we gradually heat up a diatomic gas, increasingly higher vibration levels are gradually excited by the kinetic impacts until the molecules are finally "thermally" dissociated. This can be done because the selection rules which hold for optical transitions do not apply to the change of energy states due to collisions. The reason for this is that the conservation of momentum and energy can be fulfilled in a dissociation by collision, because the process results in *three* particles—the colliding particle and the two atoms of the dissociated molecule.

On the other hand, a photodissociation is possible if an electron is simultaneously excited according to the equation

$$AB + hc\bar{\nu}_c \rightarrow A + B^* \tag{6-22}$$

We explain this process, and the determination of the dissociation energies from the spectrum, with the aid of Fig. 181. By absorption of light from the vibration ground state $v'' = 0$ of the molecule, transitions can occur (for the case illustrated) into the discrete states and into the continuous region of the upper electron state. Thus we find a band sequence converging toward a limit, with a dissociation continuum beyond this shortwave limit. The wave number of the longwave limit of the continuum corresponds to the energy L (Fig. 181). It seems reasonable to assume, and is in agreement with experience, that in a photodissociation of the type described the excited molecule decomposes not into two normal atoms but, in general, into a normal and an excited atom (e.g., $A + B^*$ or $A^* + B$), since in the process of separation of the nuclei the excited electron remains in its excited state. According to Fig. 181 we can arrive at the final state (normal + excited atom, $A + B^*$) in two ways from the initial state (normal molecule AB). Firstly, we can excite the molecule (electron transition) by transferring the excitation energy E_m of the molecule and then, by adding vibration energy of the amount D', dissociate the excited molecule into a normal and an excited atom ($L = E_m + D'$). Here D' is the dissociation energy of the excited electron state of the molecule. Since the molecular excitation energy E_m is equal to the wave number of the 0,0 band, and thus is obtained from the spectrum, and L is the observable convergence limit of the band sequence, $D' = L - E_m$ can be determined directly. In the second way we can, theoretically, first add the vibration energy D'' to the molecule, thus dissociating it into normal atoms, and then, by adding the atomic excitation energy E_a, excite one atom. We then have $L = D'' + E_a$ and arrive at the same final state, one normal and one excited atom ($A + B^*$).

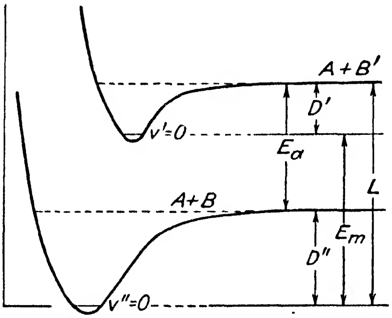


FIG. 181. Determination of the dissociation energy of a diatomic molecule in its ground state (D'') and an excited state (D'). E_m = electron excitation energy of the nonvibrating molecule; E_a = excitation energy of the excited atom B, originating from the dissociation of the excited molecule; L = energy of the limit of the band sequence $v'' = 0 \rightarrow v'$, to be taken from the band spectrum or to be extrapolated.

From this consideration it follows that *we obtain the most interesting dissociation energy of the ground state,* D'', *if we subtract from the energy value of the band convergence limit* L *the excitation energy* E_a *of the excited dissociation product (atom* B*).* The quantity E_a can be determined in many cases. If the excited atom returns to the ground state, it emits a line corresponding to the energy E_a and we get E_a directly from wavelength measurements. In the case of the oxygen molecule, for example, the value of the dissociation energy D'' of O_2 was long in doubt because E_a was unknown. Finally the analysis of the O atomic spectrum allowed E_a, and thus the dissociation energy of O_2 to be established.

Only in a few cases (e.g., in the molecular spectra of the halogens and the ultraviolet spectra of H_2 and O_2), do the potential curves have such favorable relative positions that the band limits L themselves appear in the spectra. Usually we find only a shorter or longer band sequence whose convergence limit has to be determined by extrapolation. From the band formula (6-18) we get for the dissociation energy D

$$D = \frac{hc\omega_0}{4x_0} \tag{6-23}$$

Thus we need only to find the value for the energy of the vibration quantum $hc\omega_0$ of the ground state, best taken from the empirical band formula (6-20), and that of the anharmonicity constant x_0 in order to be able to compute the dissociation energy (method of Birge and Sponer). However, reliable values are obtained from this formula only if the molecule under consideration can be represented satisfactorily by the equations for the anharmonic oscillator [(6-13) and (6-14)]. In numerous cases in which the cubic term must be used to represent the vibration, the above formulas give incorrect values of D (usually too large). In these cases it is better to extrapolate graphically, by plotting the vibration energy E_v versus the vibration quanta $\omega_{v+\frac{1}{2}}$. Then the extrapolated energy value for $\omega_{v+\frac{1}{2}} \rightarrow 0$ is the extrapolated D value. The degree of accuracy of the extrapolation of the dissociation energy must be checked for each individual case. Also for this check, the graphical method is better suited than the analytical formula (6-23).

b. *Predissociation*

Closely related to the dissociation process is the interesting phenomenon of predissociation which is, in addition to its basic interest, of high importance for photochemistry. In several spectra of diatomic molecules (e.g., S_2) and especially in those of numerous polyatomic molecules, a region of diffuse bands is observed in a sequence of normal sharp bands. Toward shorter wavelengths, the diffuse bands usually (but not always) merge into the dissociation continuum. This phenomenon was discovered by Henri and named *predissociation*. Chemical investigations proved that dissociation of the molecule occurs when the molecules are irradiated by wavelengths of the diffuse bands. Dissociation thus occurs before the convergence limit of the bands, i.e., by longer wave light than corresponds to the dissociation continuum.

As in the case of the phenomenon of preionization discussed on page 186, predissociation is caused by nonradiative transitions from quantized energy states into a region of continuous energy, which in the case of predissociation corresponds to that of the dissociated molecule. To explain predissociation, Fig. 182 shows the potential curves and vibration level diagrams of two excited molecular states. We assume that a transition from a lower state, not shown, to the state a' is possible.

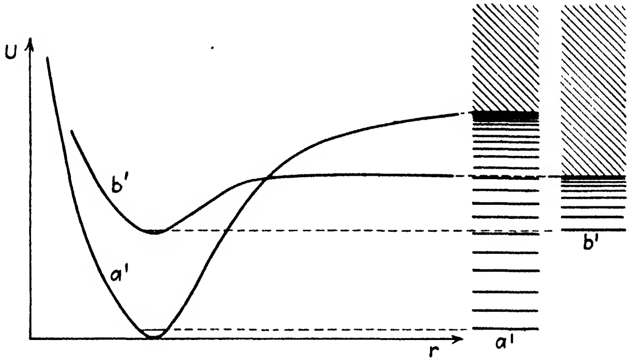


FIG. 182. Potential curves and corresponding vibration term sequences leading to predissociation of a diatomic molecule.

Then in the absence of state b' we would observe a sequence of absorption bands which corresponds to transitions from a lower state to all a' states beyond which, after the dissociation limit of a' is reached, the dissociation continuum of a' follows. Now we consider the influence of the disturbing molecular state b', whose dissociation energy lies *below* that of a'. If certain conditions (selection rules) between the electron quantum numbers of a' and b' are fulfilled, there exists for the discrete a' levels lying above the dissociation limit of b' the possibility of nonradiative transitions (without gain or loss of energy) into the continuous energy region of b', and thus of dissociation before the dissociation limit of a' itself is reached. Since the Franck-Condon principle is valid also for transition without emission or absorption of radiation, these transitions from a' into the continuum of b' occur with high probability only in the region of the intersection of a' and b', while for higher states the probability of predissociation again decreases.

As in the case of preionization (page 186), the possibility of transitions into the continuum implies a considerable reduction of the life-time of the discrete states of a' and thus, according to the uncertainty relation, an increase in the width of these energy states and the lines resulting from their combination. In the region of predissociation this line width usually becomes so large that the different rotation lines of a band merge into one another, thus accounting for the diffuse appearance of the dissociation bands. Predissociation is observed only in absorption, because, as a result of the nonradiative transitions into the continuous energy region, the radiation probability of predissociating states is extremely small.

For the photochemist this section contains some very important results: *Irradiation by those wavelengths which correspond to sharp absorption bands of the irradiated molecule produce excited, and thus reactive, molecules. Irradiation by those wavelengths, on the other hand, which correspond to continuous absorption spectra or diffuse absorption bands (which appear diffuse only because of the small dispersion of the spectroscope), produces dissociation of the irradiated molecules.* Thus the relation between the continuous absorption spectra of molecules and their decomposition into atoms is explained.

c. *The Processes of Formation of Molecules from Atoms*

What do we know about the reverse process, the formation of (diatomic) molecules from their atoms? In this case the situation is similar to the combination of an ion and an electron to form a neutral atom as discussed on page 105. If two atoms which can form a molecule do collide, they form a molecule only if the binding energy released in the collision (which of course is numerically equal to the dissociation energy) is somehow carried away. This can be done by a third atom. In that case we have the formation of a molecule in a triple collision of atoms and can write the process as an equation,

$$A + B + C \rightarrow AB + C_{\text{fast}} \tag{6-24}$$

We now consider whether this molecular recombination can also occur if *two* atoms collide and the binding energy is emitted as a photon. In order to understand this problem correctly, we must first know the difference between atomic (homopolar) and ionic (heteropolar) molecules. By *atomic molecules* we mean those which decompose into normal atoms when dissociated in the ground state (e.g., H_2, O_2, CO, etc). *Ionic* or *heteropolar molecules*, on the other hand, are those molecules which (like NaCl) dissociate from the ground state into ions (Na^+ and Cl^-).

In the usual case of atomic molecules the colliding atoms A and B "run along" the potential curve of the normal molecule (see Fig. 180). *The formation of a molecule, under emission of the binding energy and the* *relative kinetic energies of A and B, is in this case just as impossible as the reverse process, the optical dissociation without electron excitation.* The reason is that transitions from the continuous energy region to the vibrational ground state of the *same* electron state do not occur, because their transition probability, according to the Franck-Condon principle, is practically zero (Fig. 180). Two normal atoms which can form a molecule thus cannot recombine in a two-particle collision, but only in a triple collision in which the third partner, which can be the wall of a glass tube, for example, takes over the excess energy.

The situation is different in the case of an ionic molecule (e.g., NaCl). Here the colliding atoms are in a different electron state (curve V' of Fig. 183) than the ground state V'' which is formed by two ions. Consequently, a recombination in a two-particle collision is possible. In this case, the radiation probability (and with it the recombination probability) is determined by the probability of the electron transition $AB \rightarrow A^+B^-$ and by the Franck-Condon principle. In the case of the alkali halides, for example, transitions to the normal ground state can occur, according to the Franck-Condon principle, from the reversal points between F' and G' of the colliding normal atoms A and B. In this case the energy is radiated away as an emission continuum whose wavelength limits are given by the transition arrows $F'F''$ and $G'G''$, so that we have the two-particle collision recombination process

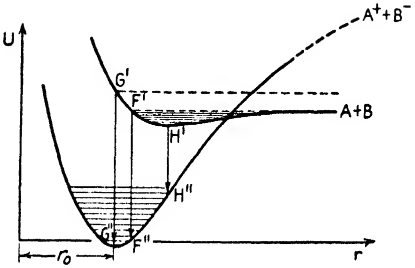


FIG. 183. Set of potential curves for explaining the formation of an ionic molecule in a collision of two atoms upon emission of the recombination continuum.

$$A + B \rightarrow A^+B^- + hc\bar{\nu}_c \tag{6-25}$$

Consequently, *just as photodissociation under continuous absorption is possible only if there is a simultaneous change of the electron configuration, the formation of molecules in a two-particle collision under continuous emission of the binding and kinetic energies occurs only if there is a simultaneous change in the electron configuration.* Such an electron transition is also connected with the radiative recombination of normal and excited atoms in two-particle collisions, according to the equation

$$A + B^* \rightarrow AB + hc\bar{\nu}_c \tag{6-26}$$

In a highly dissociated and excited gas, for example in a discharge plasma, collisions between normal and excited atoms are sufficiently frequent. These take place "along" the curve of the excited molecular state, Fig. 181. As in the case mentioned above, in collisions the kinetic energy of the colliding atoms is converted into potential energy so that a point above the dissociation limit on the left side of the potential curve is reached. By transitions from this branch of the potential curve of the excited molecule to the ground state, the colliding atoms A and B^* recombine under emission of continuous radiation.

Thus, the observation of such emission continua of molecules furnishes evidence for the process of radiative recombination, and from it the probability can be determined as a function of the velocity of the colliding atoms. Emission continua, which correspond to the recombination process (6-25), are found especially in the case of alkali-halide vapors, whereas continua corresponding to radiative recombination (6-26) have been observed for the halogens and tellurium.

6-7. Limitations of the Molecule Concept. Van der Waals Molecules and Collision Pairs

We have so far used the term "molecule" as if it were a self-evident concept, but now we must refine our knowledge and proceed to a higher point of view. By a molecule we usually mean, without giving it too much consideration, a system of two or more atoms or groups of atoms whose potential energy for a definite nuclear configuration is at a minimum, and which is, to a certain extent, independent of its environment. The latter means that (at least at the instant of observation) the interaction within the system is great compared to that between the system and its environment. Physically such a molecule is, as we know, characterized by the forces acting between the atoms, i.e., by the change of potential with the internuclear distance (potential curve) and by the value of its total energy. The energy is measured upward (positive) from the ground state of the normal molecule or (as we do here) from the zero point of the free atoms. In this case the binding energy of the stable molecular states is negative, whereas the kinetic energy of the free, dissociated states of the molecule is positive.

We now can distinguish between three types of molecules, depending on the shape of their potential curves, i.e., their binding forces (lower part of Fig. 184). We consider first the normal, strongly bound molecules which were almost exclusively treated in the last sections. Their characteristic features are a pronounced potential minimum at a relatively small internuclear distance ($r_0 \sim 1$A), and a relatively large dissociation energy of 1 to 5 ev (Fig. 184, curve a). The second type of molecule is called a "van der Waals molecule" for reasons which will become clear below. It is characterized by the potential curve b, with a shallow minimum at large internuclear distance ($r_0 \sim 3$ to 5A), and by a usually very small dissociation energy (order of magnitude of 0.01 to 0.1 ev in the ground state). These molecules consist of atoms which are not able to form a normal, strongly bound molecule by rearranging their electron shells. However, there exists between them some attraction due to mutual polarization and related interaction forces of the second order. These second-order forces imply a possibility of formation of loosely bound molecules. The same types of interatomic and intermolecular attractive forces which produce these weak bonds are also responsible for the deviations in the behavior of a real from an ideal gas. Since these deviations are described by the van der Waals equation of state, we call those molecules which are bound by these forces van der Waals molecules. Now if we consider as the third type of molecule an atomic system which is characterized by the potential curve *c*, we see that this differs from a van der Waals molecule only in that it cannot have negative energy values at all. The repulsive forces always exceed the attractive forces; in every other respect such a system of atoms behaves just as an arbitrary real molecule in a free state of positive energy. It is represented by two atoms in the instant of the collision, if only repulsive forces act between them. Such a system is called a *collision pair*. We introduce it here as the limiting case of the molecule, although it lacks a potential minimum and thus does not fit the customary definition of a molecule given above. Evidently a sharp line cannot be drawn between the van der Waals molecule and the collision pair. Furthermore,

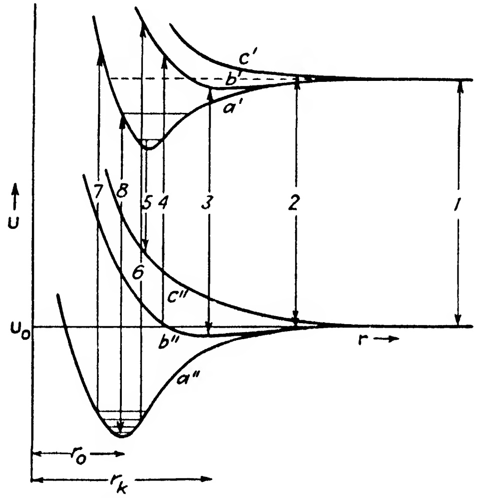


FIG. 184. Potential-curve diagram for the electron ground state and an excited electron state of a diatomic molecule with transitions indicating all possible discrete and continuous molecular spectra.

a collision pair can emit and absorb spectra just as other molecules. Finally, a pure repulsive potential curve characterizing a collision pair occasionally also occurs in the case of excited states of regular molecules, as we know, for instance, from Fig. 119.

The best known example of a van der Waals type molecule is the

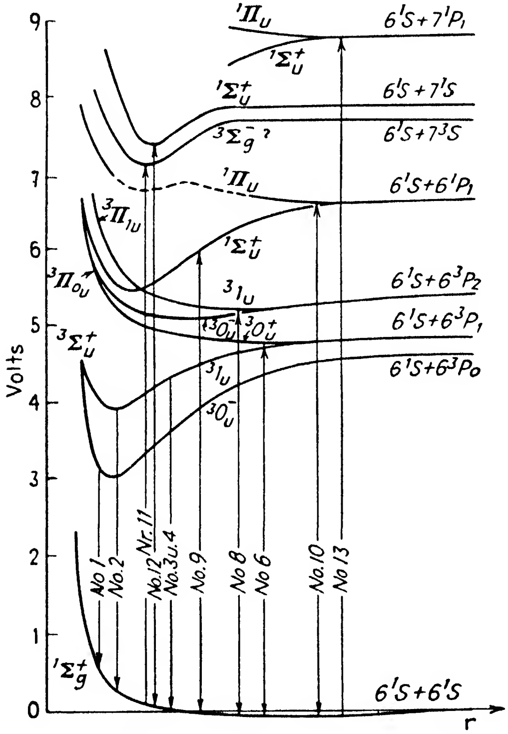


Fig. 185. Potential-curve diagram of the Hg_2 van der Waals molecule, determined from the study of the diffuse bands and continua. (*By Mrozowski and the author.*)

Hg_2 molecule consisting of two Hg atoms. All types of binding which we have mentioned above occur among its excited energy states. Figure 185 shows its potential curve diagram which is the product of many years of patient research of a large number of investigators. The He_2 molecule consisting of two helium atoms, whose existence was established spectroscopically, is another example of a van der Waals molecule, but only in the ground state. Pronounced potential minima are characteristic for the excited states of He_2 and give rise to the emission of numerous He_2 bands, which were analyzed, especially by Weizel. By transitions from these stable excited states to the unstable ground state, an emission continuum lying in the far ultraviolet is emitted.

A noticeable van der Waals attraction also exists between two O_2 molecules and, according to an investigation of the author, is responsible for the existence of the double molecule $O_2—O_2 = (O_2)_2$. It consists of two normal oxygen molecules and, besides other spectra, absorbs the continuous bands which are responsible for the blue color of compressed as well as liquid oxygen.

Van der Waals forces, whose potential corresponds to curve b in Fig. 184, play an important role in other phenomena also. We mention here only the large number of cohesion and adhesion phenomena which are produced by them, as well as the secondary forces by which smaller and larger molecules are bound together in certain crystals and colloidal micelles which play an important role in biology. They also are, at least partly, responsible for the process of polymerization of gigantic molecular complexes from single molecules (e.g., the caoutchouc molecule from about 2,000 isoprene molecules). We point out, finally, that the loose binding of two molecules effected by such van der Waals forces seems to be important for the initial stage of many chemical reactions, because it is often a prerequisite for a sufficiently intense interaction of the molecules. The often very casual treatment of the van der Waals forces and the spectra resulting from them (also the variation of an atomic spectrum by the van der Waals binding of the atom to a large complex, such as a crystal or a wall) does not seem to be justified, since a careful investigation of these phenomena may produce many interesting results.

After this extension of our knowledge about the concept of molecules and the potential curves of the different types of molecules, we return to the discrete as well as continuous spectra which result from the shape and mutual positions of the potential curves of the combining molecular states and, conversely, can be used to determine them. For this purpose we consider the schematic diagram Fig. 184 in which are shown as asymptotes to the right the two families of curves of the ground state and an excited state of an atom. The potential curves originate from bringing an additional normal atom into interaction with the first one, this being in its different states. The different possible types of transitions are represented by arrows. However, we have to consider that the transitions naturally occur also from all neighboring points of the potential curve under consideration.

Whereas the transition number 1 between the undisturbed atomic states produces a sharp spectral line, a broadened atomic line results from transition number 2. It is evident that the amount and character of the broadening depends upon the average interatomic distance (function of the density of the gas), on the temperature, and the shape of the two potential curves. With these few remarks, we have considered the process of collision broadening, already discussed on page 187, from an entirely new point of view. We considered the emitting or absorbing and the colliding atom together as a molecule (or collision pair) and thus conceived of collision broadening as a limiting case of a molecular spectrum. It is now evident that it is only a small step from the simple broadened atomic line to the spectra of the weakly bound van der Waals molecule represented by transitions 3 and 4. They consist of narrow continuous or apparently continuous bands which are closely related to the corresponding atomic line (transition number 1). The emission and absorption of mercury vapor at pressures which are not too low provide numerous examples of this type of band associated with atomic lines. We have already referred to the identification of these bands as spectra of a Hg_2 van der Waals molecule. The transitions 5 to 8 occur in the region of nuclear separation in which the electron clouds of the two atoms penetrate each other considerably and thus cause major changes of the energy compared to the unperturbed atomic states. They represent, therefore, molecular spectra in the more restricted and customary sense. Whereas the normal electron band spectra correspond to transitions 8 between the discrete states of both potential curves, there exists an extended emission continuum due to transition 5 from the minimum of an excited molecular state to the repulsion curve of a lower state. The best known examples of this are the continuum of the helium molecule in the extreme ultraviolet, which was mentioned on page 77, and the hydrogen molecular continuum extending from the green to the far ultraviolet. The latter is emitted by any glow discharge in dry hydrogen at several millimeters pressure. In conclusion, transition 7 from the discrete states of the lower potential curve to the continuous energy region of an upper curve accounts for an extended absorption continuum. The best known examples for this are the absorption continua of the halogen molecules I_2, Br_2, Cl_2, IBr, etc., which lie in the visible or near ultraviolet.

This discussion reveals how many important conclusions with respect to emission and absorption spectra can be drawn from the potential curve diagram. It is worth while, therefore, to become thoroughly familiar with the concept of the potential curve and its consequences.

6-8. Molecular Rotation and the Determination of Moments of Inertia and Internuclear Distances from the Rotation Structure of the Spectra of Diatomic Molecules

After we have become acquainted with the numerous phenomena which are related to molecular vibration and dissociation, we shall discuss now molecular rotation and the conclusions which may be drawn from investigations of the rotational phenomena in molecular spectra.

a. Rotation Term Diagrams and Infrared Rotation Spectra

The simplest model of the rigid molecule rotating about the axis of greatest moment of inertia, the rotator with free axis, was discussed wave-mechanically on page 218. We found there that only the following sequence of discrete energy values is possible for the rotator:

$$E_{\text{rot}} = hcBJ(J+1) \qquad J = 0, 1, 2, \ldots \qquad \text{(6-27)}$$

where

$$B = \frac{h}{8\pi^2 cI} \qquad \text{(6-28)}$$

is the rotation constant and

$$I = \mu r^2 = \frac{m_1 m_2}{m_1 + m_2} r^2 \qquad \text{(6-29)}$$

again is the moment of inertia of the molecule about the axis of rotation. J is the rotation quantum number belonging to the total angular momentum $\mathbf{J}$ of the rotating molecule.

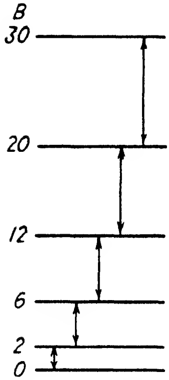


FIG. 186. Possible transitions in the rotation-level diagram of the electron and vibration ground state of a diatomic molecule (rotation band).

The rotation term diagram shown in Fig. 186 follows from (6-27). For optical transitions between these rotation states the selection rule

$$\Delta J = \pm 1 \qquad \text{(6-30)}$$

which follows from the correspondence principle (page 164) or from wave mechanics, applies. Thus each rotation state can combine under emission or absorption of radiation only with its two neighboring states. From this selection rule and Fig. 186 it follows that the rotation spectrum of a diatomic molecule lying in the far infrared consists of a sequence of equidistant spectral lines (rotation lines) with the wave numbers $2B$, $4B$, $6B$, ... The distance between two consecutive lines thus should be $2B$. Pure rotation spectra in good agreement with this theory were first found by Czerny in the case of the HCl molecule. *Just as in the case of the vibration spectrum (page 369), the rotation spectrum can be observed only if a change of the electric moment is associated with the rotation (dipole radiation). Thus the centers of the electric charges must not coincide with the center of mass. This condition is fulfilled only for the unsymmetrical molecules such as* HI, *but not for* H_2, Cl_2, O_2, *etc. Consequently, the latter do not have easily observable rotation or rotation vibration bands.*

If the rotation spectrum is observable, then the rotation constant B can be taken from it immediately. The moment of inertia of the molecule, and thus one of the most important molecular constants, can be computed from B according to Eq. (6-28). Moreover, if the masses m_1 and m_2 of the constituent atoms of a molecule are known (and this will always be the case if we know the molecule responsible for the spectrum), then the internuclear distance of the molecule follows immediately from the moment of inertia by means of Eq. (6-29). The internuclear distances determined in this way are within the order of magnitude of 1 A, which is in agreement with our knowledge of the dimensions of atoms and molecules. Table 14 presents the most important data for a number of diatomic molecules, as determined from their spectra. If the rotation and vibration are optically inactive, as in the case of molecules consisting of identical atoms, then B, and with it I and r_0, can be determined, though not quite as easily, from the rotational structure of the electronic bands, or from the rotational Raman effect, page 352.

Table 14. Internuclear Distances, Moments of Inertia, Fundamental Vibrational Quanta, and Dissociation Energies of Some Important Molecules in the Ground State

Molecule	Internuclear distance, r_0 A	Moment of inertia I, 10^{-40} gm cm²	Fundamental vibrational quantum, cm⁻¹	Dissociation energy	
				Electron volt	kcal/mole
H_2	0.77	0.47	4390	4.46	103
O_2	1.20	19.1	1580	5.09	118
N_2	1.09	13.8	2360	7.35?	170
S_2	1.60	67.7	727	4.45	103
Cl_2	1.98	113.5	565	2.47	57
Br_2	2.28	342	324	1.96	45
J_2	2.66	741	214	1.53	35
CO	1.13	15.0	2169	9.6?	220?
NO	1.15	16.3	1907	5.3	122
HCl	1.27	2.60	2989	4.40	102

b. The Rotation Vibration Spectrum

From the discussion of the pure rotation spectrum we proceed to the rotation vibration spectrum which is produced by changes in the rotational *and* vibrational states of the molecule (without a change in the configuration of the electrons). Two sequences of rotational terms, belonging to the vibration levels $v = 0$ and $v = 1$, are shown in Fig. 187. Corresponding to the selection rule (6-30), the resulting rotation vibration band consists of two "branches" which belong to the transitions $\Delta J = +1$ and $\Delta J = -1$. They are shown to the left and right in

Fig. 187. The line sequence belonging to $\Delta J = +1$ extends from the zero line (not shown) toward shorter wavelengths and is called the positive or R branch. The sequence belonging to $\Delta J = -1$ extends toward the longer wavelengths and is called the negative or P branch.

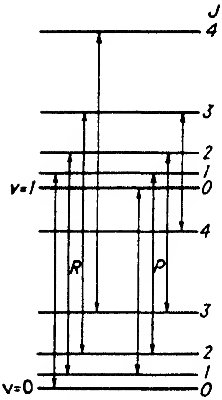


Fig. 187. Energy-level diagram for explaining a rotation vibration band of a diatomic molecule.

If we assume as a rough approximation that the moment of inertia of the molecule is the same for the upper as for the lower vibrational state in spite of the difference in vibration, then the separations of the rotation levels are equal in both vibration states. This case is presented in Fig. 187. Then once more we have equidistant lines separated by the constant distance $2B$ (Fig. 188). The zero line, which cannot appear because the transition $\Delta J = 0$ is forbidden, can be easily determined. Its frequency gives the energy of the pure vibrational transition. For every vibrational transition which is possible according to Figs. 177 and 178, we expect a band of the kind described. This corresponds roughly to the spectroscopic results. A small deviation from the assumed constancy of the distances between consecutive lines, however, is evidence that, first, the moment of inertia is not constant and independent of the vibrational state and that, second, the rotation and vibration of the molecule are not independent. The energy values of the pure rotation and vibration thus cannot simply be added, but the interaction of vibration and rotation must be taken into account by introducing a mixed term containing v and J in the expression for the molecular energy. If this is done, complete agreement between theory and spectroscopic results is achieved.

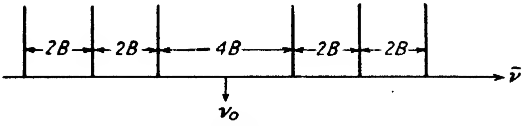


Fig. 188. Schematic representation of a rotation vibration band.

c. The Rotation Structure of Normal Electron Bands

We now take up the most complicated case, that of band spectra resulting from simultaneous changes in the electron configuration, vibration, and rotation of a molecule. These bands occur in the visible and ultraviolet spectral regions. As a result of the participation of an electron transition in the absorption and emission, there is always a change in the electric moment of the molecule. Electron band spectra thus are emitted and absorbed also by symmetrical molecules such as H_2, O_2, etc. These electron transitions also make it possible for transitions (except for the $\Sigma \rightarrow \Sigma$ transitions, see below) to occur without a change in the rotational quantum number. Thus, instead of (6-30), the rotational selection rule for electron band spectra is

$$\Delta J = 0 \text{ or } \pm 1 \tag{6-31}$$

with the limitation that transitions between states with zero rotation ($0 \longleftrightarrow 0$) are forbidden since there is no change of the electric moment connected with them. Consequently, a zero or Q branch for $\Delta J = 0$ is to be added to the branches of a band corresponding to $\Delta J = \pm 1$ (positive or R branch and negative or P branch). The zero branch can now appear, because the electron transition causes a change of the electric moment, even without a change of the rotational state of the molecule. On the other hand, no change in the electron configuration is connected with a $\Sigma \rightarrow \Sigma$ transition (page 361); consequently no zero branch can appear in this case, because now the electric moment of the molecule cannot change without a change of the rotational state.

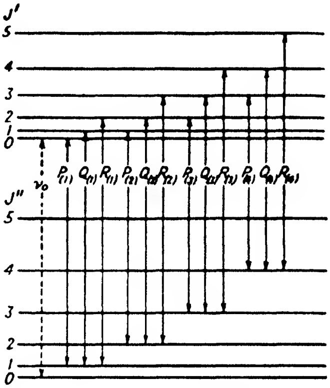


FIG. 189. Transitions between the rotation term sequences of two different electron states of a diatomic molecule. This explains the P, Q, and R branch normally present in a band connected with an electron transition.

We know from the discussion of the potential curves that the internuclear distance, the moment of inertia, and the binding forces of a molecule may be *greatly* changed by an electron transition. In general, therefore, the sequence of the vibration states as well as that of the rotational states are different in the upper and lower electron states. Accordingly, two different rotational term sequences are shown in Fig. 189, which may be thought of as belonging to any two vibration levels of any two different electron states. The energy difference of the two zero-rotation levels is the wave number of the zero line (forbidden), for which the vibration formula (6-20) applies. It is equal to the sum of the electron and vibration transitions $v' - v''$ and is written $\bar{\nu}_0(v',v'')$. The individual rotational lines of a band are designated, as in Fig. 189, by the J number of the lower state (J''). Thus the $R(3)$ line of a band is the transition $J' \longleftrightarrow J'' = 4 \longleftrightarrow 3$. From Eq. (6-27) we have the formula for the wave numbers of the rotational lines of a band

$$\bar{\nu} = \bar{\nu}_0(v',v'') + B_{v'}J'(J'+1) - B_{v''}J''(J''+1) \qquad (6\text{-}32)$$

in which, from the selection rule (6-31), J'' must be equal to J' or $J' \pm 1$. These three possible changes in the rotational quantum numbers correspond to the three branches of a band,

$$J' = \begin{cases} J'' - 1 & P \text{ branch} \\ J'' & Q \text{ branch} \\ J'' + 1 & R \text{ branch} \end{cases} \qquad (6\text{-}33)$$

which are shown in Fig. 189.

The best representation of the structure of an electron band as determined by (6-32) and (6-33) is its so-called Fortrat diagram. If, as in Fig. 190, the wave numbers of the band lines in reciprocal centimeters or wavelengths are plotted on the abscissa, and the rotational quantum numbers J'' are plotted on the ordinate, the result is a parabola for each of the three branches of the band. This can be easily ascertained by measuring the transition arrows in Fig. 189. The Fortrat diagram thus provides a means of separating the lines of the different branches and rotational quantum numbers which, in the observed spectrum, appear

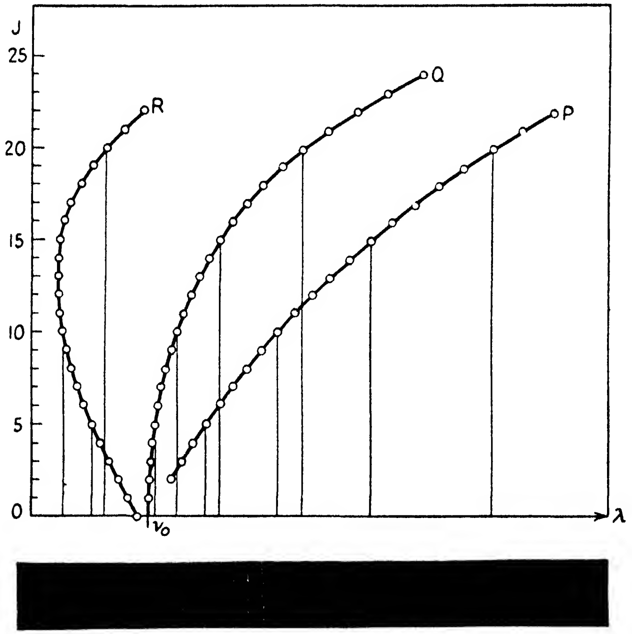


FIG. 190. Fortrat diagram of a band with three branches. Below is the whole spectral band originating from the superposition of the lines of the three branches. (*After Fermi.*)

mixed in a confusing way. Conversely, we can get from the Fortrat diagram a picture of the spectrum (in wave numbers) by projecting on the $\bar{\nu}$ axis (abscissa) the points in the diagram corresponding to the different lines. This has been done in Fig. 190, where the picture of the whole band is shown underneath the diagram. It is especially clear from the diagram that the band edge is actually not due to a convergence of the band lines, as is the series limit in an atomic spectrum (page 80), but is a more or less accidental phenomenon. It depends, just as the interval between it and the physically significant zero line (for the pure electron and vibration transition), on the intervals between the rotation levels of the upper and lower states. If the P branch forms an edge, as in our Fig. 192, we say that the band is *shaded toward the violet;* in the case of an edge formed by the R branch (Fig. 190) it is *shaded toward the red.* It follows from the relation between the Fortrat diagram and the transitions in the term diagram that the shading depends on whether the internuclear distance (and with it the moment of inertia) is greater or smaller in the upper state than it is in the lower. We thus have the following relations:

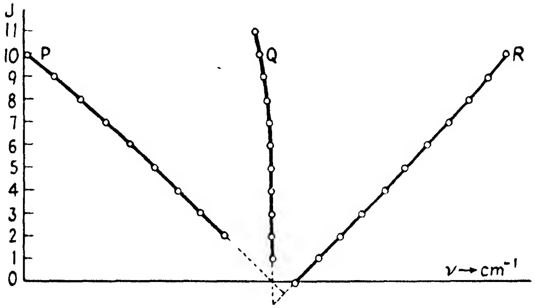


FIG. 191. Fortrat diagram of a band without edge, due to identical internuclear distances in the upper and lower electron state. (*After Weizel.*)

$$\left.\begin{array}{ll}\text{Red-shaded band } (R \text{ edge}): & r_0' > r_0''; \; I' > I''; \; B' < B'' \\ \text{Violet-shaded band } (P \text{ edge}): & r_0' < r_0''; \; I' < I''; \; B' > B''\end{array}\right\} \quad (6\text{-}34)$$

For equal values of internuclear distances and moments of inertia in the two combining states, we have no edge structure and the Q branch shrinks together to form a single line as shown approximately in Fig. 191 which is an example of an actually measured band.

Thus *a short glance at a band spectrum just to see the shading of the bands allows us to draw important conclusions about the combining molecular states, whereas an exact analysis provides us with the absolute values of the rotation constants, moments of inertia, and internuclear distances.*

d. The Influence of Electron Transitions on the Rotation Structure

In our treatment of the rotation structure of the electron bands we have not, so far, considered the fact that the total momentum **J** of the molecule which controls the spacing of the rotation terms results from the vectorial addition of the angular momentum of the rotating nuclei and an angular momentum which itself is composed, according to page 361, of the angular momentum of the electron shells about the internuclear axis **Λ**, and the resultant spin **S**. These relationships become especially complicated by the fact that the rotation of the whole molecule produces a new magnetic field which competes, with respect to the orientation of the spin of the electron shell, with the field in the direction of the internuclear axis. Therefore, the coupling between the angular momenta of the molecular rotation **K**, of the resulting orbital momentum in the direction of the internuclear axis **Λ**, and of the resultant spin of the electron shells, **S**, changes with increasing rotation, depending on the strength of the internuclear field. The different possible cases of coupling for this interaction of electronic motion and rotation have been analyzed, especially by Hund. These different cases of coupling cause differences in the rotation structure of the electron bands, such as the failure of the appearance of the zero line and of certain neighboring lines, as well as deviations from the parabolic branch curves in the Fortrat diagram. We must pass over all these details of the band structure and can only mention that, as a result of the higher multiplicity of many-electron molecules, multiple branches of bands can also appear. An investigation of all these details, however, is of interest because it permits us to determine empirically the quantum numbers Λ or Ω of the electron configuration in the upper and lower states. The complete band analysis thus provides all possible statements about a molecule. The consistent agreement of even the finest details of the often very complicated spectra with the theory is the best proof for the correctness of our theoretical concept of the molecules and their behavior.

e. The Influence of Nuclear Spin on the Rotation Structure of Symmetrical Molecules. Ortho- and Parahydrogen

In conclusion we shall briefly consider the influence of the nuclear spin on the rotation structure of diatomic molecules with identical nuclei (homonuclear molecules), upon which an interesting effect, especially in the case of molecular hydrogen, depends. We saw (pages 134 and 233) in discussing the He atom and the H_2 molecule that, because of the exchange possibility of the two identical electrons, considering also their electron spin, the term diagrams of these two-electron systems split up into two term systems, a singlet system and triplet system, which do not intercombine. *In the case of* H_2 *and the other molecules consisting of identical atoms we now have, in addition to the possibility of exchange of electrons, the exchange possibility of the identical nuclei. This again causes, according to the formalism described on page 233, the existence of two term systems which do not combine with each other.* The wave-mechanical treatment showed that the rotation terms (in the electron singlet system as in the electron triplet system of H_2) belong alternatingly to one and to the other of the two term systems which in this case are called even and odd. In the case of H_2, page 268, we have two nuclear spin moments each of $\frac{1}{2}(h/2\pi)$, which can have the same direction (analogous to the triplet system) or opposite directions (analogous to the singlet system). Because of the small interaction between the nuclear spin and the motion of the molecule, the rotation terms which belong to the three possible orientations of the parallel spin moments do not split up. The even and odd terms thus can be distinguished only by their statistical weight, i.e., the rotation band lines corresponding to them are distinguished by their different intensities (in this case 1:3). Actually, as was first found by Mecke, the band spectra of all molecules formed from identical atoms (but, for example, not that of the slightly unsymmetrical $N^{14}N^{15}$ molecules) show these *alternating intensities of consecutive band lines* which depend on the nuclear spin. In agreement with the theory, the intensity ratio in the case of H_2 is 1:3. This agreement of theory and experiment, and the resulting possibility of determining the spin momentum of the nuclei of molecules from measurements in band spectra, which supplement in a welcome way the nuclear spin determinations by magnetic methods (page 268) and by hyperfine-structure studies in atomic spectra, is an excellent example of the interrelationship of the various fields of atomic physics.

In the case of the hydrogen molecule, whose rotation states are widely separated because of its small moment of inertia, the existence of the two rotation term systems which do not combine with one another leads to an interesting effect. Because this effect is in contradiction to the classical theory, it can be regarded as another proof for the correctness of quantum physics. By continuously decreasing the temperature, not all H_2 molecules can gradually enter the rotationless state $J = 0$, as is expected classically. Only the molecules with parallel nuclear spin can enter this state while those with antiparallel spin, because of the Pauli principle, must remain in the next higher, $J = 1$, state, which cannot combine with the $J = 0$ state, regardless of how much we try to cool down the hydrogen. This fact becomes evident in a well-known anomaly of the specific heat of hydrogen at low temperatures. Thus molecular hydrogen behaves as if it consisted of two different modifications which Bonhoeffer and Harteck called ortho- and parahydrogen. It must be realized, however, that in contrast to the phenomenon of helium, which depends on the effect of the electron spin, the difference between ortho- and parahydrogen depends on the nuclear spin of the atoms which form the molecule, or, more accurately, on the exchange possibility of the two identical nuclei. A transformation of orthohydrogen (antiparallel nuclear spin momenta, $J = 1, 3, 5, \ldots$) into parahydrogen (parallel nuclear spin momenta, $J = 0, 2, 4, \ldots$) would require a change of the direction of one of the nuclear spin moments, which normally is as forbidden for nuclei as it is for electrons (page 147). By special treatment of H_2 under high pressure and by adsorption to cooled charcoal, Bonhoeffer and Harteck and, independently, Eucken and Hiller were able to produce pure parahydrogen. They found, as was expected, that the rotation lines corresponding to $J = 1, 3, 5, \ldots$ were then missing.

6-9. Band Intensities and Temperature Determination by Band Spectroscopy

Now that we have learned about the structure of the spectra of diatomic molecules and its relation to molecular constitution, we shall briefly consider the problem of the intensity of band spectra. This is important, first, because line structure *and* intensity distribution are necessary to give a complete description of a spectrum and, second, because the measurement of the intensity distribution in band spectra has led to the development of an important method for determining high temperatures.

According to page 166, the intensity of a spectrum is determined by the transition probability between the two combining states, and by the number of molecules occupying the initial state of the transition in question, its so-called occupation number. The transition probability, and with it the intensity of the entire band system, is determined by the probability of the electron transitions, computed according to page 216. The transition probability for the individual bands, on the other hand, follows from the Franck-Condon principle (page 370). Within each branch of a band, finally, the transition probability is constant, whereas the relative intensity of the different branches can be computed from J and the electron quantum numbers according to formulas of Hönl and London. A definite statement about the occupation number of the initial state can be made from Maxwell's distribution law, but only for the case of thermal equilibrium, i.e., in general for the absorption of molecules. For the case of emission, on the other hand, a statement can be made only if the excitation is due exclusively to the thermal velocity. For the intensity of a spectral line resulting from a transition between states of energies E_1 and E_2 we have

$$I_{E_{12}} = C_{1,2} g_1 e^{-E_1/kT} \tag{6-35}$$

where $C_{1,2}$ is a constant containing the transition probability between states 1 and 2, g_1 is the statistical weight of state E_1 (see page 166), and T is the absolute temperature of the absorbing or emitting molecules, i.e., of the initial state.

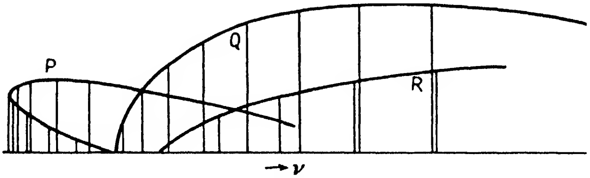


FIG. 192. Intensity distribution in the three branches of a normal band. (*After Jevons.*)

From the general formula (6-35) we obtain the intensity distribution in a rotation band if we substitute for E_1 the rotation energy

$$E_1 = hcBJ(J + 1) \tag{6-36}$$

and for g_1 the statistical weight of a rotation state

$$g_1 = 2J + 1 \tag{6-37}$$

Equation (6-37) follows from the fact that each rotational state in a diatomic molecule is $2J$-fold degenerate as a result of the equal influence of the two rotation axes which are perpendicular to the nuclear axis. With this we have

$$I_J = C(2J + 1)e^{-\frac{hcBJ(J+1)}{kT}} \tag{6-38}$$

Equation (6-38) describes the intensities of the rotation lines of a band as a function of the J values of the initial state. Figure 192 shows the intensity distribution of the rotation lines of a band computed according to (6-38). The intensity of a line is here represented by its length.

According to (6-38) the intensity distribution in a band is temperature-dependent. Thus by measuring the maximum of the intensity in a branch of the band, i.e., the J value of the line of highest intensity, the temperature of the emitting or absorbing gas can be determined. This method has been applied successfully in measuring the temperatures in electric arcs, especially in the outer zones. However, in discharges resulting from nonthermal exciting collisions, Eq. (6-38) is not applicable, and attempts to apply it lead to gross errors. Furthermore, by collisions of the second kind with energy exchange (page 104), the occupation of rotation states can be greatly changed, so that all conditions must be carefully checked before band-spectroscopic temperature measurements can be considered reliable.

If the condition of thermal equilibrium is fulfilled, then the same method described above can be used to compute how, at a given temperature, molecules are distributed among the *vibration states* of the initial electronic state. By applying the Franck-Condon principle, page 370, the intensity distribution of the bands within a band system then can be determined. In this case we have to substitute in (6-35)

$$E_1 = hc\omega(v + \tfrac{1}{2}) \tag{6-39}$$

Since the vibration states of diatomic molecules are not degenerate, the statistical weight of all vibration states is unity. Thus we have

$$I_v = Ce^{-\frac{hc\omega(v+\frac{1}{2})}{kT}} \tag{6-40}$$

Because of the difficulty of numerically computing the transition probability contained in C according to the Franck-Condon principle, Eq. (6-40) is less suited for the *absolute* determination of temperature from the relative intensities of the different bands of a system. However, the determination of *changes* of the temperature from the corresponding changes of the relative intensities of the various bands is easily possible, since in this case C remains constant. This vibration method has already been used with success in investigating temperatures, especially of electric arcs, and we may be certain that both methods of measuring temperatures by band spectroscopy will be further developed and will play an important role in the development of the physics of high temperatures.

6-10. Isotope Measurement by Means of Molecular Spectra

We have already mentioned in the discussion of the phenomenon of isotopes that the detection of isotopes as well as the determination of their masses and relative abundances is possible by optical spectroscopic methods just as by mass-spectroscopic methods. We have treated on page 181 the influence of isotopes on the hyperfine structure of line spectra; now we shall discuss briefly the isotope effect on band spectra.

Since the different isotopes of an atom differ only in their masses, the incorporation of different isotopes in the same molecule (e.g., Li^6H and Li^7H) affects the band spectra through the corresponding change in the moment of inertia (6-29) of the molecule. According to Eqs. (6-27) and (6-28), the separation of the rotational levels is inversely proportional to the moment of inertia I of the molecule, while according to Eqs. (6-11) and (6-12) the spacing of the vibration levels is inversely proportional to the square root of the moment of inertia I. The results of these two mass influences on the spectra are the *rotational* and the *vibrational isotope effect*.

We now consider two diatomic molecules, one of which may have the atomic masses m_1 and m_2; in the second molecule the mass m_1 is replaced by the mass of an isotope, $m_1 + \Delta m$. In computing, by means of the formulas on page 388 for the rotation, the wave-number difference of two equal rotation lines of these two molecules, we find

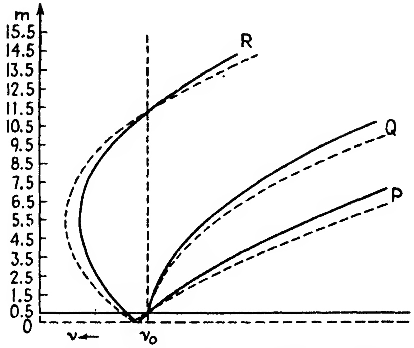


FIG. 193. Rotational isotope effect in an electron band. The branches belonging to the two isotopic molecules are indicated as solid and dotted curves, respectively. (*After Mecke.*)

$$\Delta\bar{\nu}_r = \frac{m_2\,\Delta m}{(m_1 + m_2)(m_1 + \Delta m)}\,\bar{\nu}_r \tag{6-41}$$

where $\bar{\nu}_r$ is the wave-number distance of the rotation line from the zero line of the band. From the vibration formulas on page 365 we find, for the wave-number difference of two band edges of the two isotopic molecules, the formula

$$\Delta\bar{\nu}_s = \frac{m_2\,\Delta m}{2(m_1 + m_2)(m_1 + \Delta m)}\,\bar{\nu}_s \tag{6-42}$$

in which $\bar{\nu}_s$ is the distance of the band edge under consideration from the (0,0) band of the band system. Figure 193 shows the rotational isotope effect represented in a Fortrat diagram.

Equations (6-41) and (6-42) show that the isotopic splitting of the lines increases linearly with the distance from the zero line or the (0,0) band, respectively. By measuring the spacings $\Delta\bar{\nu}_r$ or $\Delta\bar{\nu}_s$ and the distances $\bar{\nu}_r$ or $\bar{\nu}_s$, respectively, the mass ratio $m_1/(m_1 + \Delta m)$ of the isotopic atoms in the molecule can thus be determined. Because of the high accuracy of spectroscopic wavelength measurements this method of determining isotope masses is very accurate. By measuring the intensity ratios of the lines belonging to the isotopic molecules, the ratio of the two kinds of molecules, and thus the relative abundance of the isotopic atoms in the molecule, can be determined very accurately.

By means of this isotope effect in band spectra, new isotopes of the elements C, N, and O were discovered for the first time. A large number of isotopes were confirmed after they had been detected by mass-spectroscopic experiments. Mass ratios and relative abundances have been measured for many new isotopes. Mass-spectroscopic, line-spectroscopic, and band-spectroscopic isotope methods thus supplement each other very nicely.

6-11. Survey of the Spectra and Structure of Polyatomic Molecules

The results of molecular research deduced from diatomic molecules as they have been presented in this chapter can be applied, at least in their fundamentals, to polyatomic molecules. However, because of the large number of possibilities for excitation and ionization, for vibration and dissociation, as well as for the rotation about the three axes associated with the different principal moments of inertia, the spectra of these molecules are very complicated. A complete analysis of the electronic, vibrational, and rotational structure, and the corresponding complete knowledge of all molecular data, therefore, is available only for some simpler polyatomic molecules (such as H_2O and CO_2); in most cases we have to be content with incomplete information. Therefore, we shall not discuss in all detail the electronic, vibrational, and rotational structures of polyatomic molecules, but shall limit ourselves to a survey of the more important phenomena which are typical of such molecules.

a. Electron Excitation and Ionization of Polyatomic Molecules

The systematics of the electron states of polyatomic molecules is closely related to that of diatomic molecules (page 359). It has been advanced especially by Mulliken in the last 10 years. The difference compared to diatomic molecules consists, in the first place, in the larger number of external electrons which can be excited and ionized. In the second place, these electrons, depending on their location in the molecule, play different roles in holding the molecule together. We may distinguish three extreme cases of this behavior.

In many molecules the near ultraviolet absorption is due to electrons which are responsible also for holding the molecule together. For this reason they are called binding electrons. Their excitation often reduces the binding to such an extent (analogous to the case of the diatomic molecule, Fig. 169, page 364) that immediate dissociation into two atom groups results. According to page 377, the corresponding absorption spectrum must then be continuous. This photodissociation, caused by the absorption of radiation by binding electrons, explains why the ultraviolet absorption spectrum of so many polyatomic molecules is continuous. An interjection is necessary at this point. One must be careful to distinguish, in the case of polyatomic molecules, between genuinely continuous and nongenuine quasi-continuous spectra. The latter are actually discrete band spectra. The reason for this is the following: Since the distance between the band lines according to page 388 is inversely proportional to the moment of inertia of the molecule, in polyatomic molecules with large moments of inertia the distance between the band lines is often so small that the rotation structure cannot be resolved with the usual spectroscopes, even with large gratings. The discrete band spectra then *appear* to be continuous. We cannot discuss here the methods used to distinguish between the quasi-continuous and the genuine continuous spectra.

In contrast to the first case of the excitation of binding electrons which often produces dissociation, is the case investigated especially by Price, of absorption of light by nonbinding electrons. As in the corresponding case, Fig. 168, page 363 of diatomic molecules, this leads to the second extreme case where the excitation can go over into ionization. In this case the molecule remains stable during the excitation or separation of the nonbinding electron; the nuclear arrangement and vibration are independent of the motion of *this* electron. Spectroscopically, Rydberg series of bands are observed in the vacuum ultraviolet onto which is joined in many cases an ionization continuum, from whose long wavelength limit the ionization energy can be determined. However, it must be remembered that, with polyatomic molecules, one can no longer speak strictly of *one* ionization energy, because the ionization energy varies greatly depending on the binding state of the particular separated electron. Only because, for some reason not yet known to us, *one* electron is preferably excited and ionized, can its ionization energy be called that of the molecule. Our molecular theory here, as so often in atomic physics, has not yet arrived at its final state. We can determine what electron is excited and in what manner, if the molecule absorbs a definite spectrum. However, we cannot yet answer the question why, in a given electron configuration of a large molecule, one particular electron and only that one electron absorbs the incident light and becomes excited.

A systematic investigation of the third group of excitable electrons, the study of chromophoric groups, can perhaps contribute to the solution of this question. In a polyatomic molecule, radiation can also be absorbed by an electron which does not belong to the shell of the whole molecule but which is strongly localized in a definite group of atoms. This may be a benzene ring or any similar atom group belonging to a large molecular complex. In this case the resulting spectrum is, in a first approximation, independent of the large complex to which the atom group (e.g., NO_2 or $N{=}N$) is attached, and only finer detail differences (wavelength displacements) permit one to draw conclusions about the state of binding of the particular group to the whole molecule. Since the absorption spectra of this type of group electrons lie mostly in the visible spectral region and thus determine the color of the material which consists of these molecules, these absorbing groups are called color producers or chromophoric groups.

Our knowledge of the possibility of exciting particular electrons, which depended exclusively on the investigation of absorption, was extended in certain cases by the investigation of fluorescence (page 354). The investigation of the emission spectra of polyatomic molecules for a long time failed to produce reliable results because these complicated structures almost always dissociated in gas discharges as a consequence of exciting electron collisions. What could be observed, therefore, was a confusing superposition of the continuous spectra of the molecule and its components. Recent work of Schüler and Woeldike seems to have opened a new approach for the investigation of the possibility of exciting particular electrons in polyatomic molecules. By using a low-current glow discharge with special precautions, they were able to excite discrete as well as continuous emission spectra of very complicated and easily dissociable polyatomic molecules. By investigating a whole series of such molecules in which either different groups were attached to the same molecule or, conversely, the same groups were attached to different molecular complexes, and also by comparing the emission spectra with the absorption and fluorescence spectra of the same molecule, they were able to establish a large number of empirical rules and relations which seem to be of high importance for understanding the interaction in polyatomic molecules. Interesting differences have been ascertained concerning molecular excitation by light (absorption) and by electron impact (excitation in a discharge). These differences will aid us in the attempt to answer the above-mentioned question why and how a particular electron is preferentially excited in a certain process. In the case of formaldehyde, acetone, and diacetyle, for example, fluorescent light is emitted from a lower energy state of the molecule than that which is excited by the absorption of light. This initial state of fluorescence is normally reached by collisions of the second kind (page 104) from the state excited by absorption. This same initial upper state of fluorescence which thus *cannot* be reached directly by *absorption*, *can* be excited directly from the ground state by *electron impact*. In other cases, as in the case of benzene and its derivatives, light *absorption* always leads to an excitation of the nonlocalized π-electrons of the benzene ring. Recent investigations have proved, however, that in the glow discharge (i.e., by *electron impact*) this emission of the C_6 ring, which corresponds to a spectrum between 2000 and 3000 A, is excited only if the mass number of the individual substituents, or that of two substituents coupled by a double bond, is smaller than 27. Since this value is of the order of magnitude of the mass of two C atoms in the ring (mass number 24), we seem to have here a theoretically interesting relation: *the electron excitation can apparently be blocked by externally attached atom groups whose mass is greater than that of the absorbing group.* Similarly, though not quite so pronounced, the mass influence seems to be effective in the case of excitation of the molecules to fluorescence (decreasing intensity of fluorescence in the series F-, Cl-, Br-benzene). What happens to the energy which has been taken up, if the emission is blocked, seems not yet clear. It does not seem to be transformed to vibration energy. If there exists some unknown storage mechanism, it would be of fundamental interest. With reference to all these questions as well as regarding the detailed behavior of the different outer electrons of larger molecules, many interesting results may be expected from this new experimental approach. In concluding, we mention the interesting observation of Schüler and Woeldike that, for example, the sharp emission bands of quinone become diffuse when phenol is mixed with the emitting vapor. Apparently the normal, undisturbed vibration and rotation of the quinone molecule are hindered by the formation of intermolecular "hydrogen bridges," again a very pictorial result.

At the end of this section (page 407) we shall return to the question, which dissociation process is associated with the different observed absorption continua of polyatomic molecules. First, however, we shall briefly discuss molecular rotation.

b. Rotational Structure and Moments of Inertia of Polyatomic Molecules

Rotational structure, which provided so simple a method for determining moments of inertia and internuclear distances of *diatomic molecules*, is (except for the special case of linear molecules which we shall discuss soon) of little value in the study of polyatomic molecules. First, the moments of inertia of even the relatively simple polyatomic molecules soon become so large that the rotation bands can no longer be resolved even with highly resolving spectrographs, because, according to page 388, the distance of consecutive band lines is inversely proportional to the corresponding moment of inertia of the molecule. Furthermore, in the most general case, a polyatomic molecule has three principal moments of inertia corresponding to the model of the unsymmetrical gyroscope, and the rotation structure rapidly becomes extremely complicated for this model. Only in the case of the H_2O molecule was Mecke able to explain almost completely the details of the rotation structure and to determine correspondingly complete molecular data. Even in the apparently simple cases in which two or even three of the moments of inertia are equal, the rotation structure is so complicated because of the interaction between rotation and vibration that only a few molecules (such as the methyl halides CH_3I, etc., by Herzberg) have been accurately investigated. Actually only the linear molecules such as CO_2, N_2O, C_2H_2, etc. (Fig 194) have a simple and clearly arranged rotational structure. In these molecules, as in the diatomic molecules, the moment of inertia about the molecular axis is zero in the first approximation (if one neglects the deformation vibration and the resulting orbital momentum of the molecular electrons), and the other two moments of inertia are equal. In this case (Fig. 194) the rotation bands are exactly equal to those of the diatomic molecules so that the formulas of page 388 can be applied. Thus the moment of inertia can easily be deduced from the spectra. However, even if the masses of the atoms forming the molecule are known, in this case the internuclear distances *cannot* be computed without additional information because different nuclear configurations can have the same moment of inertia. Nonspectroscopic methods, e.g., electron or X-ray diffraction, thus must be used to determine the internuclear distances.

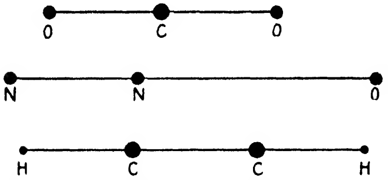


FIG. 194. Schematic representation of some linear polyatomic molecules, CO_2, N_2O, C_2H_2.

c. *Vibration and Dissociation of Polyatomic Molecules*

The problem of the vibration of an arbitrary polyatomic molecule formed of N atoms also implies many theoretical complications. In order to determine the number of possible vibrations we have to subtract from the $3N$ degrees of freedom of the N atoms the three degrees of freedom of translation and in general three degrees of freedom for the rotation of the whole molecule, so that there are $3N - 6$ degrees of freedom of vibration. Consequently, $3N - 6$ fundamental vibration frequencies also have to be taken from the spectrum (infrared and Raman). From these, the actual vibrations of the molecule originate (just as Lissajous figures of mechanical vibrations) by superposition with amplitudes corresponding to the different excitation states. These $3N - 6$ simple (in the first approximation, harmonic) vibrations from which all actual vibrations of the molecule are formed, are called the *normal vibrations.* The six normal vibrations of a plane four-atomic molecule such as formaldehyde H_2CO are shown, as an example, in Fig. 195. The sixth normal vibration is perpendicular to the plane of the molecule in which the other five vibrations lie.

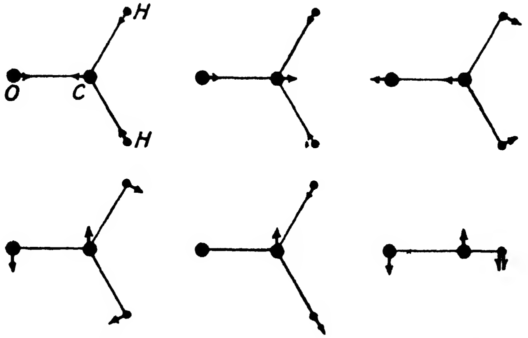


FIG. 195. The normal vibrations of the formaldehyde molecule H_2CO. (*After Mecke.*)

The fundamental difficulty in investigating the vibration of polyatomic molecules lies in the fact that, in contrast to the diatomic molecules, the spectroscopic determination of the fundamental frequencies of the normal vibrations is *not* sufficient to draw conclusions about the forces between the different atoms and atom groups. One can easily see, for example, from Fig. 195, that the number of spiral and leaf springs one must think of in order to meet the requirements of the normal vibrations shown in Fig. 195, is greater than the six normal vibrations whose frequencies can be determined from the spectrum. The number of force constants is, for f normal vibrations, equal to $f(f + 1)/2$. Consequently, *the determination of the normal vibrations from the spectrum is* not *sufficient to establish a molecular model. In addition, the geometry of the vibrations must be established, and the correct spectroscopically determined fundamental vibration frequencies must be attributed to the different normal vibration forms.* The determination of the vibration forms can be made by geometrical considerations in the case of simple molecules as soon as the nuclear configuration has been established. The vibration isotope effect, page 399, can be used to aid in establishing the form of vibration of complicated molecules. If an isotope of a different mass is substituted for any atom of a molecule, the change in the vibration spectrum produced by this substitution is the greater, the more the substituted atom participates in the vibration. By systematical substitutions the different forms of vibrations and their associated frequencies can thus be determined. Another aid is the fact established by Mecke, that the vibrations along the line joining two nuclei, the "valence vibrations" (e.g., the first normal vibration of Fig. 195), always have a higher frequency than those associated with the angular changes, the "deformation vibrations," such as the fourth and sixth normal vibration shown in Fig. 195. Thus *the difficulty in the vibrational analysis of polyatomic molecules lies not in determining the fundamental vibration frequencies, which are furnished to a large extent by Raman spectroscopy, but in attributing these frequencies to the normal vibrations of the vibration forms.*

It is self-evident that the vibrations of polyatomic molcules are anharmonic, and thus our treatment of page 367*f*. is valid also for the individual normal vibrations. In particular, each valence vibration can be represented by a potential curve such as Fig. 172, and this representation is of value in explaining the dissociation processes which we shall discuss below. From the anharmonicity of the vibration it follows, as in the case of diatomic molecules, that, in addition to the transitions between neighboring states, larger vibrational transitions also occur in the spectrum. However, these appear with smaller intensity. Moreover, it follows from the vibration anharmonicity that there is a mutual influence of the different vibrations which is manifested in the appearance of *combination vibrations*, i.e., the superposition of different normal vibrations. Obviously, this leads to an increased complication of the vibration spectrum. This situation has an analogue in acoustics where the analysis of a composite sound phenomenon, as from an organ pipe, offers many difficulties. Just as in the spectra of diatomic molecules, the frequencies of all those vibrations which cause a change of the electric moment of the molecule occur in the emission and absorption spectra of polyatomic molecules. The frequencies of those vibrations which change the polarizability of the molecule can be taken from the Raman spectrum. Again infrared and Raman spectra supplement each other in supplying the desired information, so that almost all molecular vibrations can be determined experimentally.

We distinguished above, in discussing the motion of electrons in polyatomic molecules, between localized electrons associated with a definite atom or atom group, and nonlocalized electrons associated with the shell of the whole molecule. In an analogous way, we can distinguish between *localized* and *nonlocalized vibrations* according to whether the energy of the vibration quantum is dependent upon the nature and the state of motion of its environment in the molecule or not. For example, according to Mecke, the vibration of two doubly bound C atoms in all hydrocarbons has the same fundamental vibration quantum ω_0 of about 1,600 cm^{-1} and so it can be termed strongly localized, whereas the $C—H$ vibration varies from 2,750 to about 3,150 cm^{-1} with increasing binding value of the C atom and consequently is somewhat dependent upon the environment, thus it is not localized. Such empirical rules, determined from investigations of a series of homologous molecules, are of great value for the understanding of polyatomic molecules.

Of much greater importance for polyatomic than for diatomic molecules is the interaction between vibration and rotation. It is evident from the vibration forms such as those in Fig. 195 that this vibration will be more or less distorted by molecular rotation, depending on the particular type of excited vibration. This is analogous to the perturbation of the vibration of a Foucault pendulum by the Coriolis force due to the earth's rotation. Finally there are certain cases when a molecular vibration can go directly over into a rotation. This is possible if, for example, a molecular group such as CH_3 is either free to rotate with respect to the rest of the molecule or is held in its stable position by such weak forces that it can describe slow rotatory vibrations about its rest position. Occasionally these can go over into a real rotation. Thus the clear distinction between vibration and rotation, which is characteristic for the case of diatomic molecules, can get completely lost in the case of some polyatomic molecules. We cannot discuss here the correspondingly complicated spectral phenomena.

After this survey of the vibrations in polyatomic molecules we finish our discussion by briefly considering the question of dissociation processes. As in the case of diatomic molecules, dissociation by absorption of continuous radiation due to excessive vibration is possible (page 377), but it is limited naturally to the overexcitation of valence vibrations (breaking of bonds). This clearly emphasizes the significance of valence vibrations as compared to deformative vibrations. *Because of the large number of possible vibrations, there is a large number of possibilities for the molecule to dissociate.* For this same reason, there appears much more frequently than in the case of diatomic molecules a superposition of discrete and continuous vibration term systems which, according to page 388, can lead to predissociation. In both cases, consequently, the large number of possibilities makes it often very difficult to attribute a particular absorption continuum or a wavelength of predissociation in the spectrum of a polyatomic molecule to a definite dissociation process.

In favorable cases, photochemical data can be used to compute approximate dissociation energies for the different possibilities of dissociation, and then to attribute them to the observed continua (or their maxima). This was first done by Schumacher, Stieger, and the author when they associated the longwave maxima of the pure continuous Cl_2O absorption spectrum with its decomposition into $ClO + Cl$ with and without excitation of the dissociation products, whereas they attributed the shortwave maximum to a photodissociation of the molecule into the three atoms $Cl + Cl + O$. This interpretation was confirmed by the simultaneous investigation of the photochemical decomposition. Upon irradiation by light of the different wavelengths, a different quantum yield was found which agreed with that expected for the different processes. Similar investigations also seem promising for other molecules. In certain cases, furthermore, a direct chemical detection of dissociation products may be possible and could be an aid in checking the correctness of an assumed decomposition process. Occasionally one can go farther, as was first done by Wieland in interpreting the continuous absorption spectrum of the metallic halides of the type HgI_2. He was able to identify the excited dissociation product by its emission and thus confirm the particular dissociation process. Finally, by investigating the absorption spectra of an entire series of homologous molecules and by substituting other atoms at various places in the molecule, the decomposition processes can be clarified and associated with their continuous spectra or regions of predissociation.

In general we have to understand that, with polyatomic molecules, the value of the dissociation energy determined from the long wavelength limit of an absorption continuum (or the wavelength where predissociation begins) indicates only an upper limit. Because of the coupling between the molecular vibrations, frequently, in the process of dissociation by the absorption of light, some of the energy is used to excite other vibrations.

To summarize: The investigation of the spectra of polyatomic molecules will give us answers about the excitation, ionization, and dissociation processes and the amounts of energy required for these processes. In the simpler cases it also permits us to determine the possibilities of vibration and the force constants as well as the moments of inertia. For all complicated molecules, however, the determination of the nuclear configuration and the binding forces is possible only if the spectroscopic methods are supplemented by the other methods of molecular physics mentioned on page 348, especially by electron and X-ray diffraction.

6-12. The Physical Explanation of Chemical Binding

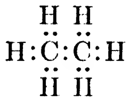
In this chapter we have discussed the most important facts about the structure and properties of molecules. In doing this, we did not consider the difficult problem, which atomic forces account for the binding of a number of atoms, either like or unlike, in a molecule. This is really the fundamental problem of chemistry, because otherwise chemistry has to accept simply as a sort of miracle the fact that there is an H_2, but no H_3, molecule, a CO and a CO_2, but no CO_3, an H_2SO_4, but no HSO, etc. The theoretical explanation of the periodic table (page 168) was the first major achievement of atomic physics in basic chemistry. The physical explanation of chemical binding, achieved by the methods of quantum mechanics in 1927, completed the successful attempt to understand the basic features of chemistry from the properties of the atoms.

Partial success had been obtained, as early as 1916, by Kossel, who succeeded in explaining the binding in heteropolar or ionic molecules (page 381), such as NaCl, as an effect of electrostatic forces. The Bohr theory, just developed at that time, explained the unique behavior of the chemically inactive noble gases from their completed (saturated) electron shells. Atoms with one or a few external electrons outside of closed shells (such as alkali or alkaline-earth atoms) are electropositive, because they can attain noble-gas configuration by release of these external electrons. The halogen atoms, on the other hand, are electronegative, because they can complete their noble-gas shell by receiving an additional electron. The binding of an alkali-halide molecule, according to Kossel, is based on the following: A Na atom, for instance, gives up its external electron to a Cl atom, thus producing a Na^+ and a Cl^- ion. These ions, which are very stable because of their ideal noble-gas configuration, attract each other electrostatically and thus form a polar molecule.

Also polyatomic molecules can be explained by this effect of ionic binding. Of the combining atoms, those from the left side of the periodic table always appear positively charged, those from the right side negatively charged. Atoms from the central groups of the table can appear as positive or negative ions depending on whether they combine with atoms which are farther to the right or farther to the left in the periodic table. The molecules PH_3 and PCl_5 are examples of this. In PH_3 the electron shell of the P atom is completed, by inclusion of the three hydrogen electrons, to form the argon noble-gas shell. It then becomes a threefold negatively charged P^{---} ion which binds the three singly positive H^+ ions. In the PCl_5 molecule, on the other hand, each of the five Cl atoms tears away an electron from the P atom, so that its electron shell is reduced to the noble-gas shell of neon and we get a fivefold positive P^{5+} ion which binds electrostatically the five Cl^- ions. In an analogous way, we may regard the C atom in CH_4 as a fourfold negatively charged, and in CCl_4 as a fourfold positively charged ion. Abegg's rule (1906), according to which the sum of the highest positive and negative valency of an atom is always 8, thus follows automatically from the possibility of completing or reducing the electron shell to that of a noble gas with eight outer electrons.

Kossel's explanation of polar binding is clearly limited to the binding of atoms whose external electron shells supplement each other to form noble-gas shells. The binding of the simplest existing molecules, such as H_2, N_2, O_2, and many others, i.e., of the so-called homopolar molecules, cannot be explained on this basis.

An attempt at an explanation was made somewhat later by G. N. Lewis with his octet theory which is still very popular with many chemists. Lewis's theory is also based on the fact that all electron shells tend to complete the noble-gas shell of eight electrons, so-called octets. By indicating each external electron (i.e., electrons outside of closed shells) by a dot, Lewis writes the ethane molecule C_2H_6 in the following way:



$$\begin{array}{ccccc} & \mathrm{H} & & \mathrm{H} & \\ \mathrm{H}: & \ddot{\underset{\cdot\cdot}{\mathrm{C}}} & : & \ddot{\underset{\cdot\cdot}{\mathrm{C}}} & :\mathrm{H} \\ & \mathrm{H} & & \mathrm{H} & \end{array}$$

By making use of the H electrons, the C atoms thus succeed in surrounding themselves by complete octets. Each electron pair, however, belongs to *two* atoms and thus is supposed to cause the binding between them. For many well-known molecules the octets cannot be complete. NO_2, for instance, must be written as

$$:\ddot{\underset{\cdot\cdot}{\mathrm{O}}}:\underset{\cdot}{\mathrm{N}}:\ddot{\underset{\cdot\cdot}{\mathrm{O}}}:$$

However, Lewis was able to show that such molecules are the less saturated, i.e., the more active chemically, the less complete their octet shells are. It is an important feature of Lewis's concept that he recognized homopolar chemical binding as an effect of electron pairs common to the two bound atoms. He could not yet give, however, an *explanation* for this binding effect of the electron pairs.

This explanation was given, on a quantum-mechanical basis, in 1927 by Heitler and London for the H_2 molecule. According to them, homopolar binding is caused by the exchange forces between the two H atoms; the nature of these forces has been treated on page 228 for this very example. If the distance between two H atoms, assumed to be far away from each other in the beginning, is gradually diminished, and the interaction of the electron clouds with the two nuclei is taken into account, a shallow minimum of the potential energy of the system, i.e., some binding of the atoms, results from purely classical electrostatic forces. However, the binding energy accounts for only a small fraction of the actually known binding energy of the H_2 molecule. We learned on page 234 that the quantum-mechanical binding is a consequence of an exchange resonance. In the unperturbed state (large distance between the two H atoms) the system is described by two eigenfunctions which belong to the same energy eigenvalue, because the two electrons are completely identical. This system is perturbed if the distance between the atoms is reduced until the electron eigenfunctions begin to overlap. The perturbation theory for this case shows that the two energy states of the same energy split up just as do the levels of the He atom in singlet and triplet terms. As a consequence of this interaction, a repulsive state of the system of two atoms is formed whose energy increases steadily with decreasing internuclear distance, and a "bound state" whose energy decreases to a minimum and then increases again with decreasing internuclear distance (Fig. 119). This potential minimum of the singlet state corresponds to the bound H_2 molecule. The numerical calculation of the binding energy and the internuclear distance of H_2 by Heitler and London did not agree too well with the data derived from the band spectra. Further refinements of the computations, however, have resulted in complete agreement with the empirical data. We have seen on page 232 that the eigenfunction of the bound state in a first approximation is obtained by *addition* of the eigenfunctions of the unperturbed system of the two atoms, whereas that of the repulsive state is obtained by subtraction [Eqs. (4-123) and (4-124)]. Explained pictorially, this result means that binding occurs between two atoms if the electron density between the nuclei (given by the norm of the eigenfunction) is large. Moreover, the bond is the stronger, the more the electron eigenfunctions overlap, because the value of the exchange integral (4-121) depends on this overlapping. We shall make use of this important result in the very pictorial explanation of the angular valences of stereochemistry.

We have seen on page 233 that, as a consequence of the Pauli exclusion principle, the spin directions of the two electrons in the stable ground state of the H_2 molecule must be antiparallel, whereas they are parallel for the repulsive state. For this reason, the magnetic saturation of the antiparallel spins is occasionally described as the essential feature of homopolar binding. The name *spin valence* is used in this connection. However, *not the spin saturation, but the sign and value of the exchange integral* [(4-121) of page 231] *are decisive for the binding.* In general, we have antiparallel spin directions in the binding case, because that leads to an electron eigenfunction which is symmetrical in the electrons, so that for antiparallel spin directions we find a maximum of the electron density between the nuclei. That spin compensation is not a decisive effect may be seen from the case of the O_2 molecule. Its $^3\Sigma$ ground state is evidence that the spin directions of the two outermost electrons are parallel. In this case the *orbital* momenta $L = 1$ of the oxygen atoms (ground state ^3P), instead of the spin momenta, compensate to form the $^3\Sigma$ ground state of the O_2 molecule. This is called *orbital valence* by Heitler.

Exchange forces, which here are considered responsible for homopolar binding, are known to us from the binding of protons and neutrons in the nucleus (page 337). Just as nuclear binding could be regarded as a consequence either of an exchange of protons and neutrons, or as a consequence of a (virtual) emission and absorption of mesons by the nuclei (page 337), we can regard the exchange forces of the molecular bond as due to an exchange of electrons, or as caused by (virtual) emission and absorption of photons.

Because of the basic difficulty of a pictorial concept of the exchange forces, it is simpler to approach the whole problem of chemical binding from the energy point of view. Chemical binding between two atoms occurs and can occur only if the system of the two atoms has, at a small internuclear distance, a state of lower potential energy than if the atoms are completely separated. Herzberg, Hund, and Mulliken have developed this approach. They consider the energy of the molecule under consideration as composed of the contributions of all individual electrons, which are thought of as being built into the fixed arrangement of the nuclei of the molecule. The problem is now whether energy has to be added or is released when a particular electron is transferred from the distant atom into the molecule, i.e., whether this electron is a "binding" or a "loosening" electron. This study is confined to only the valence electrons of the atoms which form the molecule, because the internal electron shells are saturated and therefore do not participate in the chemical bonds. A certain particular bond then can be characterized simply by the difference of the number of binding and loosening electrons. In the O_2 molecule, for instance, three of the four valence electrons of the two O atoms find places in binding electron states of low potential energy ($2p\pi^4$ and $3p\sigma^2$), whereas the fourth electron pair has to use up energy in order to reach the only available higher state ($3p\pi^2$). This electron pair thus has a loosening effect and compensates the binding effect of one of the binding electron pairs. The final result thus is that we have two uncompensated binding electron pairs which are responsible for the double bond known to the chemist for the O_2 molecule.

In an analogous way it is easy to find that two atoms with two *s* valence electrons (e.g., the atoms of the second group of the periodic table) cannot form a molecule, because only one of the electron pairs can go into a lower energy state of the molecule, whereas for the second pair only a higher energy state is available. The effects of the binding and the loosening electron pair thus compensate each other, and a strongly bound homopolar Hg_2 molecule cannot exist, in contrast to the loosely bound Hg_2 van der Waals molecule, which we discussed on page 385. It is obvious that this method of investigating the binding requires an exact knowledge of the molecular electron states, whose importance thus is nicely demonstrated to the beginner. It allows us, of course, to study theoretically all possibilities of binding between given atoms, i.e., makes possible an investigation of chemically unknown molecules.

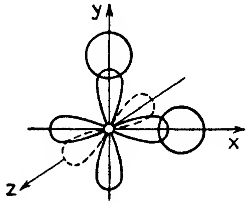


FIG. 196. Schematic representation of the binding in the H_2O molecule by *p* valences. (*After Stuart.*)

Finer details of the possible bonds and their characteristics can be discussed if we make use of the wave-mechanical "shape" of the valence electrons, i.e., their eigenfunctions according to Fig. 116 (page 227). The *p* electrons here play a particularly important role because, in contrast to the *s* electrons with their spherical symmetry, they have an axial symmetry. The three *p* electrons of the nitrogen atom, for instance, are oriented in such a way that their axes are mutually perpendicular. In the H_2O molecule, for example, the two *s* electrons of the H atoms are bound by two *p* electrons of the O atom (bond type $s—p^2—s$). As these two *p* electrons are oriented perpendicularly to each other, it is evident that H_2O cannot be a linear molecule. According to Fig. 196, the *p* electron along the *x* axis can bind only an H atom whose nucleus is also on the *x* axis, whereas the second *p* electron of the O atom, because of its orientation along the *y* axis, can bind only an H atom whose nucleus is near the *y* axis, for only then a sufficient overlapping of the electron eigenfunctions is possible. *From the wave-mechanical "shape" of the p electrons, i.e., from the* $|\psi|^2$ *distribution in space, we have thus deduced the important result that the* H_2O *molecule* cannot *be a linear molecule but should have a right angle.* In a similar way, we get, quite generally, the angular valences of stereochemistry.

The band analysis of H_2O, carried out first by Mecke, has led to an angle of somewhat over 100°, and similar deviations from right angles are found quite generally in stereochemistry. These deviations can be explained in a very pictorial way by the assumption that the atoms at the sides of the angle interact with each other and thus increase or decrease the valence angle.

An example of the binding of three s electrons by the three mutually perpendicular p electrons of another atom is the ammonia molecule, NH_3. Its pyramidal shape thus is easily understood to be a result of the binding of one H atom on each of the x, y, and z axes, and an additional attraction among the H atoms.

More complicated is the case of the tetravalent carbon atom which is of decisive importance for all organic chemistry. We shall discuss briefly the simplest organic molecule, methane, CH_4. According to Table 10 (page 178), the C atom has, outside its closed K shell, two $2s$ electrons and two $2p$ electrons. It happens, however, that these four valence electrons of the carbon atom have approximately the same energy. We speak of an approximate s-p degeneracy. By their mutual interaction, the difference between s and p electrons disappears to such an extent that the C atom has four practically equivalent valence electrons. As a consequence of general symmetry rules, the CH_4 molecule then should be, in agreement with experience, a regular tetrahedron with the C atom in its center. The particular role which the C atom plays in organic chemistry thus seems to be due to a peculiarity of its electron configuration, the approximate degeneracy of the two s- and the two p-valence electrons. On page 410 we have treated the same CH_4 molecule on the basis of the Kossel theory of polar binding. We therefore should like to use this opportunity to point out that polar and homopolar binding are not nearly such fundamentally different effects as it might appear at first sight. If, in the time average, one of the two exchanging electrons is always near the atom A, and one near the atom B, we have a regular atomic molecule and so-called homopolar binding. If, on the other hand, both electrons which are responsible for the bond are predominantly near the atom A and none near the atom B, A appears to be negatively and B positively charged. AB is then a polar molecule (such as NaCl), and we speak of polar or ionic bonds. Thus polar binding also can be understood on the basis of the exchange concept. Between these two limiting cases there exist innumerable transitions, i.e., molecules AB in which the two binding electrons are not always, but more often, near atom A than near B, so that we might speak of partial ionic and partial homopolar binding. We shall not go into any detail; we want to indicate only the fact that basically every bond can be understood by one theory, and that ionic and homopolar bonds are only the two limiting cases.

The different bonds discussed so far are designated as *localized bonds,* because every electron pair can clearly and definitely be attributed to a bond between two atoms and thus causes the bond between them. This concept of localized binding seemed to be the only possible one. However, the most important molecule of organic chemistry, the benzene ring C_6H_6, could not be explained on this basis. If one of the four valence electrons of each C atom binds one H atom, each C atom has three free valence electrons. In the benzene ring we thus should find single and double bonds alternating between the C atoms, and this actually was the opinion of its discoverer, Kekulé. This hypothesis, however, is in contradiction to the chemical and physicochemical evidence which points clearly to a completely symmetrical ring. E. Hückel has solved this problem by wave-mechanical methods. He was able to prove that each C atom is bound to its two neighboring C atoms by single bonds, and that *the remaining six valence electrons of the six C atoms are distributed uniformly over the whole benzene ring.* We thus might speak of one and one half bonds between each two C atoms. As there are no one and one half electron pairs, it is more correct to speak of one electron pair (one bond) between every two C atoms, whereas the remaining six electrons are called *"nonlocalized" valence electrons.* In our pictorial concept of electron exchange it seems quite reasonable to assume that electrons can not only be exchanged between two neighboring atoms (localized bond), but that the six nonlocalized electrons exchange their places in such a way that in the time average always one electron is near each of the six C atoms of the benzene ring. Quite a number of additional details of chemical bonds in organic molecules can also be explained by quantum mechanics, among them the rigidity of the double bond against torsion.

The atomic valence theory in its quantum-mechanical form thus is able to account for all details of the fairly complicated phenomena of chemical binding. Statements can be made also about the binding character of unknown molecules. No doubt, the mathematical difficulties of the perturbation theory as applied to large and complicated molecules become so insurmountable that in practical cases we have to confine ourselves to qualitative studies and statements. Nevertheless, the basic solution of the problem of the chemical bond is of highest importance for theoretical as well as practical chemistry. For this reason it may be especially well suited to demonstrate to the chemist the necessity of familiarizing himself with atomic physics, including the systematics of electron states and their quantum-mechanical treatment.

Literature

General Molecular Physics

Eyring, H., J. Walter, and G. E. Kimball: Quantum Chemistry, Chapman & Hall, Ltd., London, 1944.

Fermi, E.: Moleküle und Kristalle, J. A. Barth, Leipzig, 1938.

Glasstone, S.: Theoretical Chemistry, D. Van Nostrand Company, Inc., New York, 1944.

Hellmann, H.: Einführung in die Quantenchemie, Deuticke, Leipzig, 1937.

Rice, F. O., and E. Teller: The Structure of Matter, John Wiley & Sons, Inc., New York, 1949.

Stuart, H. A.: Molekülstruktur, Springer-Verlag, Berlin, 1934.

Wolf, K. L.: Theoretische Chemie, J. A. Barth, Leipzig, 1941.

Spectra (except Infrared and Raman) and Molecular Properties

Finkelnburg, W.: Kontinuierliche Spektren, Springer-Verlag, Berlin, 1938. (Reprinted in 1944 by Edwards Bros., Inc., Ann Arbor, Mich.)

Finkelnburg, W., and R. Mecke: Band Spectra, vol. 9/II, Hand- und Jahrbuch der Chemischen Physik, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1934.

Gaydon, A. G.: Dissociation Energies and Spectra of Diatomic Molecules, John Wiley & Sons, Inc., New York, 1947.

Herzberg, G.: Molecular Spectra and Molecular Structure, Prentice-Hall, Inc., New York, 1939.

Pearse, R. W. B., and A. G. Gaydon: The Identification of Molecular Spectra, Chapman & Hall, Ltd., London, 1941.

Pringsheim, P.: Fluorescence and Phosphorescence, Interscience Publishers, New York, 1949.

Sponer, H.: Molekülspektren und ihre Anwendung auf chemische Probleme, 2 vols., Springer-Verlag, Berlin, 1935–1936.

Weizel, W.: Bandenspektren, Ergänzungsband I, Handbuch der Experimentalphysik, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1931.

Infrared and Raman Spectra and Molecular Structure

Herzberg, G.: Infrared and Raman Spectra of Polyatomic Molecules, D. Van Nostrand Company, Inc., New York, 1945.

Hibben, J. H.: The Raman effect and its chemical applications, *Am. Chem. Soc. Monographs*, 1939.

Kohlrausch, K. W. F.: Der Smekal-Raman-Effekt (mit Ergänzungsband), Springer-Verlag, Berlin, 1931–1938.

Kohlrausch, K. W. F.: Ramanspektren, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1943.

Rowlins, F. I. G., and A. M. Taylor: Infrared Analysis of Molecular Structure, Cambridge University Press, London, 1929.

Schaefer, Cl., and F. Matossi: Das Ultrarote Spektrum, Springer-Verlag, Berlin, 1930.

Sutherland, G. B.: Infrared and Raman Spectra, Methuen & Co., Ltd., London, 1935.

Ta You Wu: Vibrational Spectra and Structures of Polyatomic Molecules, National University of Peking, 1939.

Valency

Kossel, W.: Valenzkräfte und Röntgenspektren, Springer-Verlag, Berlin, 1921.

Kronig, R. de L.: The Optical Basis of the Theory of Valency, The Macmillan Company, New York, 1935.

Lewis, G. N.: Valence and the Structure of Atoms and Molecules, Reinhold Publishing Corporation, New York, 1923.

Palmer, W. G.: Valency, Classical and Modern, Cambridge University Press, London, 1944.

Pauling, L.: The Nature of the Chemical Bond, 2d. ed., Cornell University Press, Ithaca, New York, 1942.

Penney, W. G.: The Quantum Theory of Valency, Methuen & Co., Ltd., London, 1935.

Rice, O. K.: Electronic Structure and Chemical Binding, McGraw-Hill Book Company, Inc., New York, 1940.

Sidgwick, N. V.: The Electronic Theory of Valency, Oxford University Press, New York, 1927.

Incidental Subjects of Molecular Physics

Debye, P.: Polar Molecules, Reinhold Publishing Corporation, New York, 1929.

Jordan, P.: Eiweissmoleküle, Wissenschaftliche Verlagsgesellschaft, Stuttgart, 1947.

Polanyi, M.: Atomic Reactions, Williams and Norgate, Ltd., London, 1932.

CHAPTER 7

ATOMIC PHYSICS OF THE LIQUID AND SOLID STATE

7-1. General Review of the Structure of the Liquid and Solid States of Matter

We now turn from the polyatomic molecules to the consideration of the very large atomic and molecular complexes, i.e., to liquids and solid bodies, though the liquid state can be treated only very briefly in this section. Because of the symmetry of the binding forces between atoms in the equilibrium state, i.e., in the minimum of potential energy, the atoms or molecules arrange themselves in a very symmetrical geometric pattern or lattice which we call a crystal. For weaker binding forces or at higher temperatures, the thermal motion of the atoms prevents them from attaining this state of equilibrium. The atoms or molecules then have a definite mobility and the regularity of the atomic pattern is more or less disturbed. This is the picture of the liquid state of matter.

The solid state, consequently, in its ideal form is identical with the solid crystalline state. In the apparently noncrystalline or amorphous structures we have either a large number of coalesced microcrystals or, as in the case of glass, an undercooled liquid which is not a genuinely solid body at all. The very few solids which are regarded as "really amorphous" at the present time, such as the "explosive" modification of antimony and one modification of selenium, apparently are not *stable* modifications of the solid state, as is evidenced by the explosibility of the amorphous antimony. The rest of this chapter will be devoted to a thorough discussion of the solid crystalline state. In this section, therefore, we shall take up briefly the problem of the structure of the liquid state, of which not too much is known at present.

From the above interpretation it is clear that in liquids, in which the average interatomic distances, because of their equivalent densities, are essentially the same as in crystals, the forces between atoms or molecules also tend to establish a crystallike, regular arrangement of the atoms. However, the crystalline structure is disturbed more or less by the thermal motion of the atoms, molecules, or ions. This explains that in liquids we find phenomena of order as well as of disorder. For the same reason, a completely random distribution of atoms cannot even be expected in compressed gases, and evidence of some degree of order in such gases has been found by the X-ray method which we shall discuss below. Although for a long time it had been assumed that the lack of a geometrically regular molecular arrangement was a characteristic of liquids, recently the indications of order have been studied in more detail. It has been concluded from this evidence that there is good reason to speak of a semicrystalline state of fluids, as O. Lehmann did so vividly 50 years ago in his "World of the Liquid Crystals."

What at first appears to be a radical distinction between the solid crystalline and the liquid state, is evidently a difference in the degree of mobility of the constituents in the crystal lattice which more or less retains its regular structure. We shall see on page 454 in the discussion of diffusion in crystals that even in the most ideal solid crystals the constituents have a certain mobility, and that a number of very important effects depend upon this mobility. This mobility of single lattice constituents (and with it the general disorder of the crystal) increases with increasing temperature until at a definite temperature the melting point is suddenly reached, above which a certain mobility is a characteristic property of *all* particles. Even in the fluid state the mobility increases more and more with increasing temperature, and this must cause a gradual disappearance of the crystalline properties. That we do not have a gradual transition from the solid to the liquid state, but a sharp melting point, is due to the fact that only the solid and the liquid state are thermodynamically stable, but not the intermediate states, where the solid crystal begins to loosen up appreciably.

Because the assertion of the semicrystalline structure of liquids is so surprising at first, we shall briefly discuss the most important evidence in favor of it.

Debye's method of scattering of X-rays by liquids provides one direct proof. If the distribution of molecules and the intermolecular spacing were completely random, the scattered intensity would decrease uniformly with increasing scattering angle. However, if there is a regular molecular arrangement, then maxima and minima are to be expected as a result of the interference of the X-rays scattered by the regularly spaced molecules. This was actually observed. If for a number of possible geometrical arrangements (i.e., liquid structures) the theoretically expected scattering curves are deduced and these compared with those obtained empirically, the molecular structure in a liquid can be determined with considerable accuracy in those cases which are not too complicated. Recently the atomic arrangement in liquid mercury has been studied in this way by Hendus who used monochromatic X-rays. At 18°C the atomic arrangement agreed essentially with that of crystallized mercury, and thus deviated considerably from the closest spherical packing which was expected formerly for the liquid state of mercury.

A no less significant evidence of the semicrystalline structure of liquids is the fact that the measured value of the atomic heat of monatomic liquids such as mercury or liquefied argon is 6 cal/mole. This is twice the value to be expected for freely mobile atoms each of whose three translational degrees of freedom contributes $R/2 = 1$ cal/mole to the atomic heat. On the other hand, the constituents of a crystal vibrate about a rest position, so that $3R/2$ cal are contributed to the atomic heat for the kinetic vibration energy and an equal amount for the potential energy, since in a harmonic vibration half of the total energy is on the average kinetic and half potential energy. Consequently the atomic heat of a solid body (if all degrees of freedom are excited) is $2(3R/2) = 6$ cal/mole. That this same value was found for monatomic liquids can only be explained by the hypothesis that their atoms vibrate in three dimensions about a rest position, except that in the case of liquids this center of vibration itself participates in a random translatory motion which is temperature dependent.

There are a number of further results of optical and electrical measurements, which we cannot discuss here, that indicate just as clearly the crystallike arrangement of the molecules in liquids. Stuart and Kast have investigated the dependence of liquid structures on the type of molecule, on the temperature, and density by model experiments. Their model liquids consisted of large numbers of differently shaped little magnets, which were shaken continuously as a substitute for the thermal motion, and in which the intermolecular forces were replaced by the forces between the small magnets which replaced the molecules. The different structures of dipole- and quadrupole-molecule liquids and quite a number of finer details can be shown very nicely in this manner.

Generally the existence of strong dipole or quadrupole moments introduces a complication in the normal liquid structure due to the formation of molecular chains (e.g., in alcohols) or molecular clusters as in the case of the so-called associated liquids whose anomalous behavior depends on the formation of these clusters. By far the most important associated liquid is water whose anomalous behavior has long been attributed to association. It had been thought earlier that water consisted of polymeric molecules of the type $(H_2O)_n$ with a constant degree of association n. However, all attempts to determine n have failed. From this failure and other evidence, the conviction has grown that water consists of molecular clusters of indefinite size. The exact structure of these molecular clusters can be determined by comparing the experimental X-ray scattering curves with the theoretically deduced patterns. By this method a tridymite-like structure was found in which, according to Fig. 197, each O atom is tetrahedrally surrounded by four protons. This particular geometric arrangement of the H_2O molecules, *i.e., this semicrystalline structure of water, is responsible for the well-known abnormal properties of water.* From this, the change in the structure and properties of water as a result of the addition of relatively few ions (or the addition of a little alcohol) can also be understood: the tridymite structure of water which depends on secondary binding forces (van der Waals forces, page 386) is greatly distorted by the electrostatic forces of the ions or by the addition of a few large foreign molecules, which at the very least influence the size of the molecular clusters. Conversely, the addition of a few H_2O molecules to pure alcohol cannot noticeably alter its chain structure. In agreement with this, hardly any change in the properties of alcohol is noticeable when it is mixed with small amounts of water. Many empirically known properties of liquids can be understood in this way from their crystal structure, i.e., from the viewpoint of atomic physics.

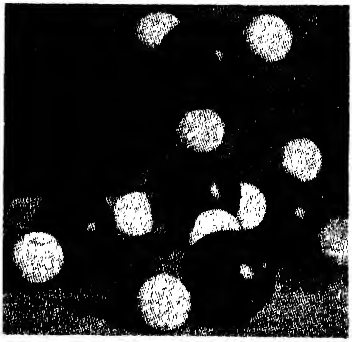


FIG. 197. The quasi-crystalline tridymite structure existing in liquid water due to association of the H_2O molecules. Black spheres = oxygen atoms, white spheres = hydrogen atoms. Each O atom is tetrahedrally surrounded by four H atoms. (*After Kortüm.*)

So far as can be concluded from the very incomplete investigations, in general, the quasi-crystalline structure of liquids is the more pronounced the greater the dipole moment or the larger and more complicated the individual molecules. Evidence for this can be derived from many remarkable phenomena which Lehmann found in his studies of large and complicated organic molecules, described in his book on liquid crystals. Most surprising among his results is the double refraction of certain liquid droplets and, in many cases, a very pronounced dichroism (different absorption in different directions of orientation). Both of these phenomena are clear evidence for an anisotropy which is characteristic of a crystal. On the other hand, the internal mobility, i.e., the change of the shape of the drops, was just as evident. Therefore, Lehmann's designation as "liquid crystals" seems correct in spite of the objection of many crystallographers. Thus we have to distinguish between the special phenomenon of liquid crystals and the general phenomenon of the solid crystals. A discussion of Lehmann's results from the standpoint of modern atomic theory, unfortunately, is still lacking in spite of the recently increasing interest in the structure of liquids.

With these remarks we close our brief discussion of the liquid state. Whenever in the following the term crystal is used, it applies exclusively to the normal, solid crystals.

7-2. Ideal and Real Crystals. Structure-sensitive and Structure-insensitive Properties of Solids

Physics of the solid state is more or less identical with crystal physics. We begin our discussion by stressing the important difference between *ideal* and *real* crystals. The ideal rock salt crystal, for example, consists of negative chlorine ions and positive sodium ions arranged in a regular geometric lattice in which each sodium ion is surrounded by six chlorine ions, and conversely. However, this ideal crystal is only a *model* and does not actually exist, first, because crystal growth does not occur quite regularly, but occasionally there is a lattice defect (e.g., a missing constituent ion) and, secondly, because the purity of the crystal material is not so high but that here and there an impurity atom is incorporated in the crystal lattice. Actually, even the best real crystals have at least one missing ion or impurity atom among 10^6 normal constituents, i.e., according to page 18, at least 10^{16} lattice defects per cubic centimeter.

Smekal first pointed to the fact that an important difference in the properties of crystals is related to this distinction between ideal and real crystals. He designated those properties of crystals which were essentially independent of lattice defects or impurity atoms as *structure-insensitive*. Among these structure-insensitive properties we may name the lattice structure, the specific heat, elasticity, thermal expansion and compressibility, the energy of formation, the principal features of optical absorption (color) and dispersion, as well as, for metallic crystals, the electronic conductivity, and finally dia- and paramagnetism. Those properties of crystals, on the other hand, which depend essentially on the number, arrangement, and type of lattice defects and impurities, are called *structure-sensitive*. Among these are especially the diffusion phenomena, the ionic and electronic conductivity in insulating crystals and semiconductors, the internal photoeffect, plasticity, and crystal strength, as well as certain finer features of optical absorption and luminescence which are of importance for the typical structure-sensitive phenomenon of phosphorescence.

The crystal defects consist, on the one hand, of lattice flaws, most of which are either holes in the lattice or lattice atoms outside their regular geometrical position (interstitial atoms), and, on the other hand, of impurity atoms in normal or abnormal lattice points. A real crystal occasionally consists of many small, almost ideal microcrystals of about 1μ diameter with defects near and on the boundaries between the individual microcrystals (layer or mosaic structure of a crystal).

In the following we shall treat first the ideal crystal and the structure-insensitive properties of solid matter. In the second part of this chapter we shall then discuss the real crystals and those phenomena and properties of solids which are caused by lattice defects and impurities.

7-3. The Crystal as a Macromolecule. Ionic, Atomic, and Molecular Lattices

We have already interpreted the crystal as a limiting case of a very large but unusually regular and symmetrical molecule. This point of view presents only one side of the picture and thus has its limitations. However, it allows us to understand many crystal phenomena without further explanation. It seems clear, for instance, that there can be unlocalized electrons in crystals, just as in the benzene ring molecules (page 415), which are only part of the crystal as a whole. Conversely, there are other electrons, just as the emitting electron of a chromophoric group of a complicated polyatomic molecule (page 402), which we can attribute to a definite group of atoms or even to a definite atom or ion of a crystal. In discussing the optical properties as well as the conductivity of crystals, this difference will prove to be very important.

A parallel can also be drawn between molecules and crystals with respect to the nature of their bonds. Thus we find that in crystals there is a classification according to the type of binding which corresponds to what in molecules we called polar, homopolar, and van der Waals binding. The *ionic crystal lattice* with simple ionic binding corresponds to the polar molecule, and it is apparent that the constituents of the ionic molecule Na^+Cl^- can form only an ionic lattice in the crystal. The alkali-halide crystals are typical representatives of this type of lattice. The *atomic crystal* corresponds to the molecule bound by homopolar forces. Here we must distinguish between two types, those with localized bonds (by electron pairs), and those with nonlocalized binding. The first group is called *valence crystals* and its best known representative is the diamond. The metals and alloys comprise the group with nonlocalized binding.

The *molecular crystal* corresponds to the van der Waals molecule (page 384). It should be called *van der Waals crystal,* because not only molecules, but also atoms can act as crystal constituents (as in the case of the solidified noble gases). If the lattice constituents are molecules (e.g., H_2, Cl_2, CO_2, CH_4), whose atoms can be bound by polar or homopolar forces, the molecules themselves are held in the lattice by the weak van der Waals forces (page 384). Accordingly, these molecular crystals (just as the noble-gas crystals) loosen up, i.e., melt at very low temperatures; in other words, they require for their decomposition only a small thermal energy. All organic compounds, for example, crystallize in molecular lattices. As a transitional type we should mention the "layer crystal" type in which (to choose at random) LiOH, CdI_2, and $CrCl_3$ crystallize. In these crystals we have a genuine atomic or ionic binding (homopolar or polar) within the crystal layer, while the individual crystal layers are held together only by van der Waals forces.

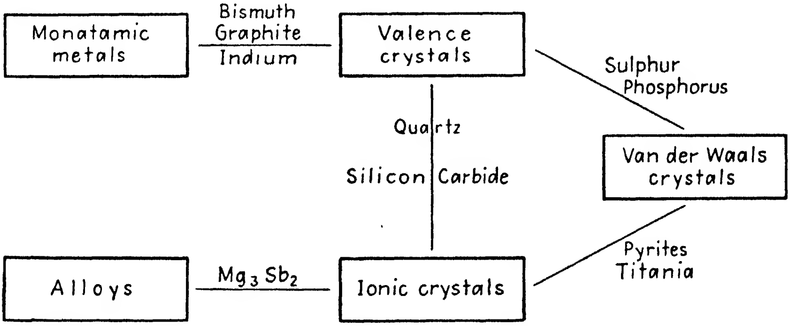


FIG. 198. Schematic diagram of the relation between the different types of solids. (*After Seitz.*)

A diagram according to Seitz (Fig. 198) presents the relation between the different types of solids in a pictorial manner. The monatomic metals have a definite relationship to the valence crystals, between which the crystals of bismuth, indium, and graphite form transitional cases. The alloys, and according to Pauling also the pure monatomic metals, have some relationship to the polar-bound ionic crystals. Evidence for this can be derived from the transitional case of certain intermetallic compounds, such as Mg_3Sb_2, with fairly well localized bonds of the ionic type. Furthermore, there are transitions between the valence and ionic crystals, just as in the analogous case of molecular binding (page 414). Thus one can consider the crystals quartz, SiO_2, and Carborundum, SiC, as valence crystals with some contribution of ionic binding. Nor are the van der Waals crystals completely isolated, since solid sulphur and phosphorus represent transitional cases between them and the valence crystals and, on the other hand, ice, pyrite, and TiO_2 form transitional cases to the ionic crystals.

Before we consider further the question of lattice binding, at least for the more important cases of ionic crystals and metals, we must briefly review the geometric arrangement and the determination of the lattice structure of a crystal.

7-4. Crystal Systems and Structure Analysis

We have already explained that because of the symmetric forces acting between the atoms or ions of a solid body (crystal) a stable equilibrium is possible only if the lattice constituents are arranged in a regular geometric pattern. By group-theoretical methods it can be proved that there is a finite but large number (230) of crystal systems which are possible arrangements of the lattice constituents in a crystal. In which one of these systems a particular material crystallizes depends upon the size of the lattice constituents and upon the magnitude and orientation of the forces acting between them. The same material can crystallize in different systems depending upon the type of binding forces; for example carbon can crystallize in the graphite or diamond lattice. The determination of the structure of a crystal, the so-called structure analysis, often is not simple; it is the task of the crystallographer. It is achieved almost exclusively by means of X-ray diffraction following the procedure of von Laue or Debye-Scherrer. We need not go into the details of these methods here. Some remarks on the basis of the method may suffice.

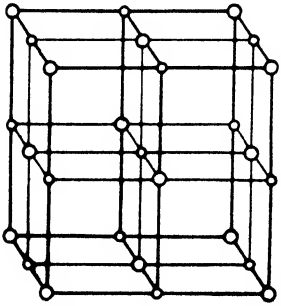


FIG. 199. Unit cell of the rock salt crystal NaCl. The Na^+ and Cl^- ions are indicated by their different size.

It follows from the periodicity of the lattice structure that for each crystal there is a smallest element, the *elementary cell*, from which the whole crystal can be built by its translatory repetition in the three coordinate directions. In the cubic NaCl crystal the elementary cell (Fig. 199) is a small cube; in most cases the form of the elementary cell is much more complicated. If a coordinate system is constructed whose axes are parallel to the edges of the crystal, then each crystal plane (and thus the whole crystal as determined by the crystal planes) can be characterized by specifying the lengths from the origin to the intersections of every crystal plane with the axes. These lengths are measured in units of the corresponding edges of the elementary cell. The system generally applied for characterizing crystal planes uses not the values from the origin to the intersections of the planes with the axes themselves, but their reciprocals which, when reduced to smallest integral numbers, are called the *Miller indices.* Thus, for example, the

(123)-face corresponds to the intersected lengths 1, $\frac{1}{2}$, $\frac{1}{3}$ or to the smallest integral numbers 6, 3, 2. The (100)-face, on the other hand, is a crystal plane parallel to the *yz* plane.

To determine the system of a given crystal, one uses the fact that when X-rays traverse a crystal, they are reflected by the atoms occupying a lattice plane in the crystal, or, in other words, they are diffracted by the lattice points. The relation between the deflection angle α (to be taken from the Laue pattern), the wavelength λ of the X-rays, and the distance d between successive lattice planes is given by the Bragg formula

$$2d \sin \alpha = n\lambda \qquad n = 1, 2, 3, \ldots \tag{7-1}$$

Moreover, we can see immediately from the Laue pattern, Fig. 200, that the symmetry of this particular crystal with respect to the direction of the incident X-rays is a threefold symmetry, i.e., a rotational periodicity of 120°. By taking Laue diagrams, with directions of incidence parallel to the different crystal axes and to the diagonals, a clear picture of the symmetry of the whole crystal can be obtained, whereas the absolute distances between the different lattice planes follow from Debye-Scherrer photographs (page 74) taken with monochromatic X-rays of known wavelength. We have indicated this crystal analysis only quite briefly because it is of no further interest to us in the atomistic treatment of the properties of solid bodies.

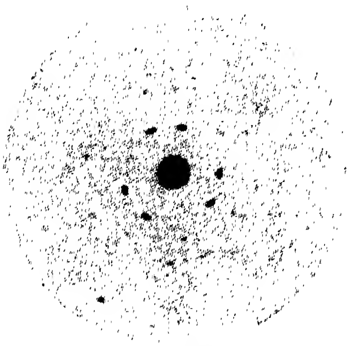


FIG. 200. Laue diagram of a crystal with three-fold symmetry.

However, the so-called *Fourier analysis*, a theoretically difficult and rather cumbersome quantitative evaluation of X-ray diffraction photographs, taking into account also the intensities of the radiation scattered in different directions, provides not only the arrangement of the crystal constituents in the crystal but also a means of determining the electron distribution effecting the binding between the atoms or ions. Figure 201 shows as an example the electron density in an NaCl crystal in which the lines indicate the places of equal electron density and the numerals, the number of electrons per cube of 1 A on a side. This method is especially important in investigating the character of the crystal binding in certain cases of transition between one type of binding and another. We have already discussed this Fourier analysis on page 348 in our discussions of the methods of research in molecular physics.

7-5. Lattice Energy, Elasticity, Compressibility, and Thermal Expansion of Ionic Crystals

After this survey of the methods of determining crystal structures we shall consider crystal binding, and the phenomena depending upon it, more carefully for ionic crystals. If we choose the model of the ideal rock salt crystal then the binding depends on the electrostatic attraction between the singly charged ions Na^+ and Cl^-. Consequently, the potential of this force of attraction between two isolated ions is

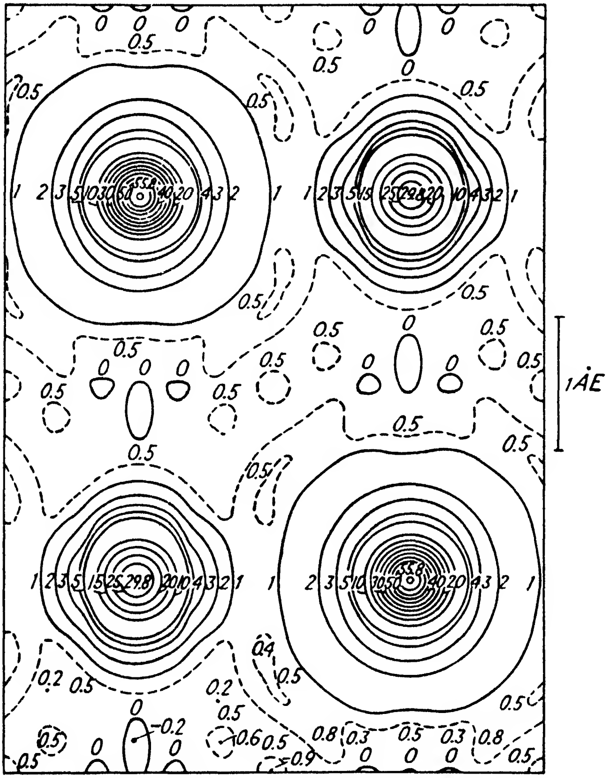


FIG. 201. Distribution of the electron density in a NaCl crystal, determined by Fourier analysis of X-ray data. The numerals indicate the electron density per cube of 1A on an edge. (*After Grimm, Brill, and coworkers.*)

$$u(r) = -\frac{e^2}{r} \tag{7-2}$$

However, the two ions are not isolated in space but are in an environment of similar positive sodium and negative chlorine ions. One can easily see that the nearest neighbors of the ion pair under consideration, because of their opposing charges, tend to reduce the binding. Quite generally, all ions of the environment will contribute to the binding energy positive or negative amounts, which decrease with increasing distance from the ion pair under consideration. The exact computations as carried out by Madelung show that the above expression for the attractive potential must be multiplied by a factor 0.29 in order to take into account the contributions of the other lattice constituents.

Since the two ions cannot penetrate each other, there must be a repulsive force in addition to the attractive force. This repulsive force can be described in a pictorial manner. Any attempt to decrease the distance between two ions, if they touch each other, requires a deformation of their shells, and this deformation requires a considerable amount of energy. For this reason, the potential of the repulsive forces increases with a high power of $1/r$. The experimental results can best be described by the ninth power of $1/r$. Thus we have for the potential of the forces acting between two ions

$$u(r) = -0.29\frac{e^2}{r} + \frac{c}{r^9} \tag{7-3}$$

where c is a constant. In order to determine its value, we make use of the fact that for the equilibrium internuclear distance r_0 the potential energy must be a minimum,

$$\left(\frac{\partial u}{\partial r}\right)_{r=r_0} = 0 \tag{7-4}$$

By performing the differentiation we get

$$c = \frac{0.29}{9}e^2r_0^8 \tag{7-5}$$

and thus

$$u(r) = -0.29e^2\left(\frac{1}{r} - \frac{1}{9}\frac{r_0^8}{r^9}\right) \tag{7-6}$$

By plotting the potential $u(r)$, just as we have done in molecular physics (page 367), we get the diagram Fig. 202. So far we have only considered the binding between a Na^+ and one of its neighboring Cl^- ions. The term *lattice potential* or *lattice energy* means the total energy per mole of the crystal, i.e., the energy required per mole to form the crystal from widely separated Na^+ and Cl^- ions. To compute the lattice energy, we have to consider that each Na^+ ion in the lattice has six neighboring Cl^- ions, each of which is bound to it by the same $u(r)$ computed above. Furthermore, 1 mole of a NaCl crystal has N_0 Na^+

ions from each of which binding forces extend to each of its six neighbors. Thus the lattice energy (referred to a normal crystal with an equilibrium distance r_0 between two consecutive constituents) is

$$U = 6N_0u = -1.74N_0e^2\frac{8}{9}\frac{1}{r_0} \tag{7-7}$$

This lattice energy is negative because it is released if the crystal is formed from ions which are assumed to exist. If we let the numerical values which refer to the specific case of an NaCl crystal and the constants be contained in a specific constant C, and if we take into consideration that r_0 is equal to one half of the lattice constant d, then the energy of all ionic crystals can be expressed by

$$U = C\frac{e^2}{d} \tag{7-8}$$

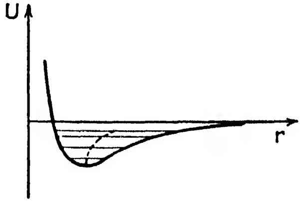


FIG. 202. Interaction potential $u(r)$ between two crystal constituents. The average internuclear distance is indicated as a function of the vibration energy.

The lattice energies computed according to this theoretical formula cannot be compared directly with experimental data, because the experimentally measured heat of formation E_F of the NaCl crystal refers not to the ions which make up the crystal but to metallic sodium and Cl_2 molecules. However, Born has shown that the lattice energy U can be computed from the measured heat of formation E_F, provided some atomic data are known or can be measured. We denote the measured heat of formation of the crystal by E_F, the heat of vaporization of sodium metal (always referred to 1 mole) by V, the dissociation energy of the chlorine molecule by D, the ionization energy of the sodium atom by E_i, and by E_a the electron affinity energy, which is set free when an electron is attached to a Cl atom. We can then write the Born cycle in the following way:

$$\left.\begin{array}{lll}
\text{Na (metal)} + \text{Cl}_2 - E_F & = & \text{NaCl (crystal)} \\
\text{Na (metal)} + V & = & \text{Na (atomic)} \\
\text{Cl}_2 \quad + D & = & 2\text{Cl} \\
\text{Na} \quad + E_i & = & \text{Na}^+ \\
\text{Cl} \quad - E_a & = & \text{Cl}^-
\end{array}\right\} \tag{7-9}$$

From this the lattice energy required to form a NaCl crystal from the Na^+ and Cl^- ions is

$$U = E_F - V - \tfrac{1}{2}D - E_i + E_a \tag{7-10}$$

Table 15 shows the good agreement of a number of lattice energies computed from the theoretical formula (7-7) with those determined from experimental data by means of the Born cycle (7-10).

Table 15. Lattice Energies of Several Alkali-halide Crystals in Cal/Mole. Comparison of the Measured and Computed Values

Alkali-halide crystal	Experimental value	Computed value
NaCl	183	182
NaBr	170	171
NaI	159	158
KCl	165	162
KBr	154	155
KI	144	144
RbCl	161	155
RbBr	151	155
RbI	141	138

From the curve of the mutual potential $u(r)$ of two lattice constituents of an ionic crystal (7-6), derived on page 428, a number of important macroscopic properties of crystals can be easily deduced or explained in terms of the atomic theory.

The compressibility is the relative decrease in volume per unit of pressure. That energy must be applied to reduce the volume, i.e., decrease the equilibrium distance r_0, can be seen from the potential curve. Moreover, it can be seen immediately that the compressibility of a crystal is the smaller, the steeper the potential curve rises from the minimum toward smaller internuclear distances. The force F necessary to reduce the distance can be computed by differentiating $u(r)$ with respect to r. Thus the work per mole required to reduce all distances from r_0 to r is

$$A = 6N_0\int_{r_0}^{r} F\,dr = \int_{r_0}^{r} p\,dv = \frac{P}{2}\,\Delta v \tag{7-11}$$

Here p is the instantaneous pressure which increases with the compression and P is the final pressure corresponding to the total change in volume, v. Here the function $p(v)$ has been approximated by a linear function. Carrying through the computation results in an expression for the compressibility,

$$K = c'\frac{d^4}{e^2} \tag{7-12}$$

The compressibility thus increases with the fourth power of the lattice constant d, a plausible result since the compressibility must be the smaller the more densely the lattice components are packed in the normal lattice.

If the compressing force is removed, the equilibrium distance r_0 is restored, since without external forces equilibrium can only exist if all distances correspond to the minimum of the potential curve. *Consequently, the crystal is elastic.* From this fact it follows that the crystal can vibrate, i.e., periodically increase and decrease the distance between the lattice points. We shall consider this important property in some detail in the section following the next.

The last important conclusion which can be drawn from the potential curve concerns the thermal expansion of crystals and, in general, of all solid bodies. The vibration energy of the crystal and thus the amplitude of the crystal vibrations increases with increasing temperature. Because of the asymmetry of the potential curve (anharmonicity of the lattice vibration) the average distance between two lattice points, shown in Fig. 202, increases with increasing amplitude of the vibration. Consequently, the volume of the crystal increases with increasing temperature. Thus *the fundamental thermal property of all known solid bodies, their expansion with increasing temperature, has a very simple explanation from the atomic theory. In the last analysis it depends upon the asymmetry of the potential curve, $u(r)$.*

We have limited ourselves in this section to a discussion of ionic crystals because their properties can be accounted for quantitatively. Qualitatively, the behavior of valence crystals, such as diamond, with respect to the potential curve and the conclusions drawn from it, such as compressibility, thermal expansion, etc., is very similar to that of the ionic crystals. Phenomenologically, valence crystals are characterized by their great hardness and good electrical insulating properties. The latter is due to the fact that electrolytic conductivity, which will be considered in detail for ionic crystals on page 457, is not possible because the constituents are atoms. Electronic conductivity, on the other hand, is impossible because, due to the strongly localized valence binding, electrons are so strongly bound that there are no conduction electrons available.

7-6. Review of the Binding Forces and Properties of Metals

Now that we have considered the properties of insulating crystals with strongly localized bonds, we shall briefly discuss the binding and the atomic properties of the other extreme of solids, the metals. They are characterized by a lack of localized binding. Such a review of the properties of metals appears to be the more appropriate as metals have to be regarded, from a practical technical standpoint, as one of the most important groups of solids. Moreover, many of the properties of metals, their modifications (different forms of lattices in different temperature regions), and their alloys were first explained by atomic physics which also threw new light on the problem of the technological treatment of metals (e.g., hardening and tempering).

In the course of this chapter we shall consider and explain the characteristic differences between insulators and metals from different points of view. At this time it is sufficient to conclude from their high electrical and thermal conductivity, that the outermost electrons of the atoms forming the metal (in general, one per atom) are not tightly bound to their atoms as they are in most insulators. On the contrary, they must have a high mobility within the metal just as do the molecules of a gas. This is why we often speak of the electron gas in the metal, a degenerate electron gas, according to page 239. These external atomic electrons which, as valence electrons, are also responsible for binding (page 409), are thus by no means localized. According to page 415, the last six valence electrons of the C_6H_6 molecule are distributed, as nonlocalized electrons, among all six C atoms of the benzene ring, and thus participate in its binding, though they do not form electron pair bonds. In a similar way the nonlocalized conduction electrons of a metal must be responsible for the binding in the metallic state (metal crystal). However, again there is no mutual saturation of electron pairs to form localized bonds. In contrast to homopolar binding, *metallic binding thus does not show the phenomenon of saturation.* Whereas two homopolarly bound H atoms cannot bind an additional H atom, in a metal each additional atom is bound with the same force independently of the number of atoms already bound in the crystal.

We know from experience that the elements on the left side of the periodic table exhibit metallic character, whereas localized binding predominates in crystals formed from the elements on the right side of the table, and it is here that we find the insulating crystals. According to Hume-Rothery, the reason for this distinction is that metals generally crystallize in a body-centered or face-centered cubic lattice or in a hexagonal, close-packed structure. On the other hand, the elements on the right side of the table belong to very different lattice types in which the coordination number (the number of surrounding atoms at the same smallest distance from one atom) is always equal to the valence of the atom under consideration. According to Fröhlich this difference depends upon the repulsive Coulomb forces between each one and all the outer electrons. Sodium and chlorine, for example, each have the valence 1. Sodium has only one outer electron, whereas chlorine, which lies on the right side of the periodic table, has seven. The binding effect of the one electron pair in the Na—Na and Cl—Cl pairs therefore is different, because in Cl—Cl we have an additional repulsion between the two times six excess electrons. Such a repulsion is not present in the case of Na—Na. Nevertheless, in the case of the diatomic Cl_2 molecule, there is a strong homopolar binding because the exchange forces (page 412) produce a unilateral charge displacement (and thus an attraction between the atoms) which compensates for the repulsion due to the other electrons. However, such an asymmetry of the electron configuration is incompatible with the high symmetry (coordination number) of the crystal lattice. Thus, because there is no Coulomb repulsion, sodium atoms can be bound by metallic binding in the Na-metal lattice, whereas for Cl atoms such a symmetrical binding is not possible. The Cl atoms can form Cl_2 molecules by unilateral deformation and reduction of their internuclear distance. These molecules can then be bound in a crystal lattice by the relatively small van der Waals forces (page 386). This contrast, which was explained here as an example for Na and Cl, *holds in general for all elements on the left or right side of the periodic table: on the left side we have metallic binding with high symmetry and coordination number, on the right complicated lattices strongly influenced by the valence and always without electrical conductivity.*

We may also mention that other important properties of all solids, such as compressibility and plasticity, are related to these differences. Both decrease in going from the left to the right, as a result of the increasing effect of the repulsive action of the outer electrons.

Metallic binding is thus characterized by the lack of valence saturation, and therefore by a high symmetry (equally strong binding of all equally distant neighbors independent of the valence of the metallic constituents). We shall try to understand the reason for this behavior of metals. Because of the high metallic density, the distance from atom to atom in the lattice is smaller with reference to the atomic radii than in most insulators. We have, therefore, a strong overlapping of the electron shells of the original atoms which form the metal lattice. This overlapping means, according to the exchange theory (page 233), that there is a high exchange probability of the outermost electrons between neighboring atoms, and thus a large exchange binding energy. However, Dehlinger and Pauling have proved that we may also look at the problem from another point of view. We may speak of a normal formation of electron pairs, and thus emphasize the relation between metallic binding and normal valence binding. However, we have then to consider these electron pair bonds between one atom and its neighbors as rotating, so that over an average of time we have again a symmetrical (nonlocalized) distribution of the electrons which cause the binding. It can be shown thermodynamically that this state with rotating bonds can be more stable at higher temperatures than the states of unilateral homopolar or polar binding. This case seems to be realized in many metal alloy lattices which have higher coordination numbers at higher temperatures than they do at lower temperatures.

Pauling has extended this picture by the additional assumption (whose correctness he was able to show thermodynamically), that ionic states also contribute to the normal metallic bonds. Thus a monovalent alkali atom in a metal occasionally should have two valence electrons and thus, as a negative ion, should be able to bind simultaneously two of its neighbors.

In general it can be said that in metals the nonselective metallic binding predominates and determines the metallic character. Nevertheless, the individual properties of the atoms, especially their spin, orbital momentum, and the valence determined by them, play a certain role. Thus, in addition to the general metallic binding, there is a certain amount of heteropolar or homopolar binding. Consequently, *a metal always crystallizes in the lattice of highest symmetry which is compatible with the special properties of the atomic lattice components.* For example, in the case of the alkalies, it follows from the diamagnetism of the metal that the spin momenta of the outer electrons of the atoms are alternatingly antiparallel. For this case, quantum-mechanical computations (by means of the exchange integral, pages 231 and 232) show that the body-centered lattice, in which each atom is surrounded by eight atoms of oppositely directed spin, is more stable than the lattice expected for pure metallic binding, in which each atom is surrounded by 12 others at equal distances, half of which have one spin direction and half the other. The energetically necessary antiparallel spin directions of the alkali electrons, which are empirically demonstrated in the diamagnetism of the alkali metals, thus are responsible for the fact that, not the lattice corresponding to pure metallic binding with a coordination number 12, but the body-centered lattice with coordination number 8, has the lower state of potential energy. Any other structure and any other spin distribution has a lower stability. The body-centered lattice which is expected theoretically has actually been found for the alkali metals by X-ray experiments.

In a similar manner the empirically established lattice structure of metals and alloys, which in turn accounts for many other properties of metals, can be explained in great detail from the point of view of atomic physics. An especially good example is presented by the different modifications of iron, i.e., the different stable lattice structures at different temperatures, which are only partly ferromagnetic. These can be understood by a detailed discussion of the effect of exchange energy (page 451). The second technically significant example of the cooperation of metallic and homopolar or polar binding is that of the alloys, i.e., mixed crystals of different metals. In general, substitutional mixed crystals are observed in alloys (if the difference in the atomic radii is not too great), in which both metals occur in arbitrary proportions. However, it follows from sudden changes of the electrical resistance as well as of the magnetic properties, that for certain rational proportions of the components (corresponding, for example, to the compounds AuCu, $AuCu_3$, FeCo, or Ni_3Fe) especially stable so-called superstructures occur which provide further evidence that the general metallic binding is superimposed by some localized binding which exhibits saturation characteristics. These localized valence forces which cause an ordered structure of alloys (e.g., Cu ions in the center and Zn ions at the corners of the elementary cubes in β-brass) are very weak forces. Evidence for this is the observation that at temperatures of only a few hundred degrees this ordered structure disappears and is replaced by a disordered lattice with a probability distribution of Cu and Zn ions.

For the sake of completeness, we mention that, besides the substitutional alloys, there exist also interstitial alloys, in which small ions like H, C, or N occupy interstitial places (see page 454). The systems palladium-hydrogen and iron-carbon are examples of such interstitial alloys.

Among the important properties of metals, to be explained by atomic physics, we finally mention ferromagnetism and supraconductivity. The former will be discussed after we have learned more about the behavior of electrons in metal crystals, whereas supraconductivity cannot as yet be satisfactorily explained in spite of some promising approaches.

7-7. Crystal Vibrations and the Determination of Their Frequencies from Infrared and Raman Spectra

Because there is a quasi-elastic binding force effective between the lattice constituents of a crystal, vibrations between them are possible just as they are between the nuclei of molecules (page 365). The assymetrical potential curve $u(r)$ (Fig. 202) is evidence that we are dealing with anharmonic vibrations. We have already used this fact in explaining the thermal expansion of solids.

We have to distinguish between the so-called *internal vibrations* and the *lattice vibrations proper* of a crystal. Internal vibrations occur particularly in molecular crystals (page 423), in which real molecular vibrations are possible within the molecules which form the lattice constituents. These molecular vibrations may be disturbed more or less by the crystal lattice, but have nothing else to do with the whole crystal.

In the lattice vibrations proper, on the other hand, the whole crystal participates. To illustrate these two types of vibrations we shall use the calcite crystal $CaCO_3$. In this crystal the lattice components are the ions Ca^{++} and $(CO_3)^{--}$. The vibrations of all ions of one kind with respect to the ions of the other kind are called the external or lattice vibrations, because the whole lattice participates in this vibration. On the other hand, there are internal vibrations within the CO_3 group (e.g., the O atoms vibrate with respect to the C atom) which have nothing to do with the calcite crystal as a whole. Therefore they occur with the same frequencies in other crystals containing the CO_3 group, such as $FeCO_3$ or $MgCO_3$.

We have to distinguish two types of lattice vibrations. In the first type, all identical lattice constituents vibrate in phase (*optical vibration*), whereas in the second type the vibrations are not in phase (so-called *acoustic vibrations*). In the first case we have, at least in ionic crystals, a vibration of all positive ions with respect to all negative ions, i.e., a vibration of the whole Na^+ lattice with respect to the Cl^- lattice. As a result of the corresponding change of the dipole moment this vibration must appear in the infrared absorption spectrum. For this reason we speak of optical vibrations. In the other extreme case the vibration progresses in the crystal as a longitudinal wave. Since in this case similar atoms do not vibrate in phase, the changes in the dipole moments largely compensate each other. Therefore these vibrations do not appear in the infrared absorption spectrum and are called acoustic vibrations since longitudinal waves of this type are known from acoustics. They are less important for crystal physics than the optical or fundamental vibrations of the lattice.

The frequency of the fundamental vibration associated with a change of the dipole moment (which may originate from the polarization of the vibrating atoms) can be taken from the infrared absorption spectrum of the crystal. The absorption coefficient of the corresponding wavelengths is very large, because all lattice ions participate in this vibration. Absorption measurements therefore are made with very thin crystal laminae. A more elegant and simpler method, although not so exact, is the reststrahlen method of Rubens. Because of the relation between absorption and reflection, the crystal reflects strongly those wavelengths which correspond to the fundamental vibration, whereas the other wavelengths are mainly transmitted by the crystal. If a light ray is reflected back and forth many times between two plates of the crystal under investigation, then all wavelengths except those of the fundamental vibration are gradually eliminated by transmission. The residual rays, or reststrahlen, remaining after many reflections thus consist essentially of the fundamental vibrations whose wavelengths can easily be measured. Occasionally it happens that, as a result of special symmetry relations which exist, for example, in fluorspar CaF_2, the fundamental vibration is not associated with a change of the total dipole moment and thus it is optically inactive. In that case it cannot be found in the absorption spectrum.

If the polarizability of the lattice constituents is changed by such optically inactive vibrations, as is almost always the case, the frequencies of the vibration under consideration can be determined by Raman effect measurements as in the corresponding case in molecular physics (page 352). This implies the advantage of working in the visible or ultraviolet region, so that optical spectrographs with their higher resolving power and dispersion can be used. Ultraviolet and Raman investigations again supplement each other so that there is no basic difficulty in determining the frequencies of the fundamental lattice vibrations.

7-8. Electron Configuration and Electron Transition Spectra in Crystals. The Energy Band Model

After the discussion of the types of solids, crystal structure, and crystal vibrations, we turn to the phenomena depending upon the electrons in crystals. We investigate first the electron configuration. The problem of the electron configuration in crystals and of the spectra resulting from electron transitions can be approached from very different points of view, all leading to the same result.

We consider first the transition from an undisturbed atom to one which is strongly disturbed by its environment (atom in a plasma, page 187), and finally to one which is incorporated in a crystal as a greatly disturbed lattice atom. By this very pictorial approach, we arrive at an essentially correct picture. Undisturbed atoms have, according to page 166, the sharp energy levels of the bound and the continuous energy regions of the free (ionized) electrons. In the spectra of plasma atoms which are disturbed by the microfields of their environment, the broadening of the higher energy states of the electrons as a result of the perturbation is already very noticeable (page 187). For the highest energy states just below the ionization limit, it can no longer be decided whether they still are to be considered as discrete energy states of the bound electrons or as continuous energy regions of the free electrons (reduction of the effective ionization potential). Finally, if we go to the lattice constituents in a crystal, we expect a different behavior, aside from the increasing magnitude of the perturbation and a broadening of the energy states, only insofar as *there is no longer a completely* *free electron. An electron released from its original atom will now move "quasi-freely" throughout the crystal in the periodic potential field of all lattice ions.* We have thus arrived, in a pictorial way, at the correct picture of the electrons in a crystal. It is evident and clear that by incorporating the atom in a crystal lattice, we do not influence the inner electrons which are responsible for X-rays. *These innermost electrons remain strongly bound to their respective nuclei; their energy states are practically undisturbed and therefore remain sharp.* Evidence for this is the sharpness of the X-ray lines emitted by metallic anticathodes (see Fig. 67, page 119). *The perturbation and thus the breadth of the energy levels of electrons increases greatly with increasing principal quantum number n. In the case of the optical energy levels, which in general are not occupied by electrons, this width reaches several electron volts, so that one can readily speak of* energy bands *of the electrons which are now already quasi-free.* We shall see that these quasi-free electrons are responsible for the electronic conductivity of the metals (page 443). For the highest energy levels the band width actually becomes so great that there is an overlapping of the different bands. We shall discuss this when we take up the theory of metallic conductivity. An energy band diagram is shown in Fig. 203.

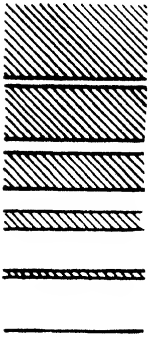


FIG. 203. Energy-band diagram of the electrons in a crystal (schematic).

It is interesting to show how one can arrive at the same result by an entirely different approach. For that reason we mention two other concepts. They are less pictorial but they have the advantage that they permit an exact computation of the behavior of the crystal electrons.

One approach is based on the concept of resonance or exchange splitting which was discussed in Chap. 4. Quantum mechanics shows (page 230), that in the case of two coupled atomic systems of equal energy their common energy state splits into two energy states, as a result of this energy resonance. The separation of these new states is the larger, the stronger the coupling between the two systems. This case exists in crystals where in principle an exchange of electrons by any two lattice constituents is possible because of their complete identity. Thus if the crystal consists of N atoms, then as a result of the possibility of exchange of each electron with the N-1 analogous electrons, each energy state of the atoms in the crystal splits up into N energy levels. We distinguish these N energy levels of every energy band by quantum numbers k. This k is closely related to the momentum of the particular electron and thus corresponds to the orbital quantum number l, introduced on page 130 for the bound atomic electrons. According to the Pauli principle, each k state can be occupied by two electrons of opposite spin. *The magnitude of the splitting, and thus the width of the band resulting from the N levels, depends on the degree of coupling, i.e. on the exchange probability of the electrons* [Eq. (4-121, page 231]. We discuss this electron exchange by means of the potential curve diagram Fig. 204. In Fig. 105 (page 187) we had represented the course of the potential near a hydrogen nucleus together with the energy states of the electron which moves in this potential field. *In a solid, because of the geometrically regular arrangement of the atoms, we have a three-dimensional periodic repetition of potential wells and potential walls,* which are indicated in Fig. 204 with the corresponding energy bands. The innermost electrons of the crystal atoms lie almost completely within the potential troughs of their ions. An exchange of these electrons with corresponding electrons of other atoms is impossible classically. Because of the tunnel effect (page 234), it is not quite impossible according to quantum mechanics, but it is very improbable because of the height of the potential wall which must be penetrated. Because of the small exchange probability, the splitting of the energy levels, and thus the breadth of the energy bands of the innermost electrons, in agreement with our first picture and with experience, is very small. Conversely, the exchange probability and the breadth of the energy bands resulting from the energy levels is very great for the quasi-free external electrons. Thus our picture is in agreement in its results with our first grossly pictorial derivation. However, it shows that *the energy bands consist theoretically of discrete energy levels and are practically continuous only because of the large number of these levels. This number per mole of the crystal equals the number of lattice atoms and thus amounts to* 6×10^{23}.

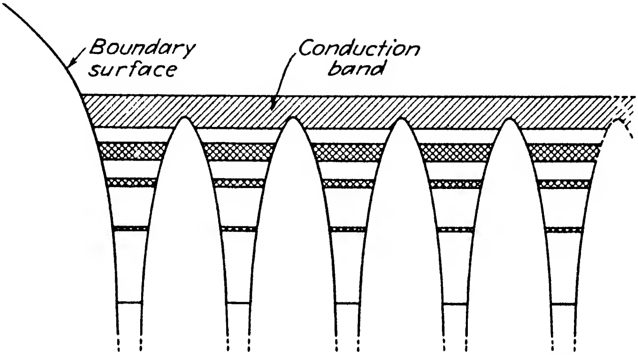


FIG. 204. Potential field in a crystal with electron states, excitation bands, and ionization (conduction) band (schematic).

In our third method of representation we start from the quasi-free electrons in the crystal and we shall show that there must be certain forbidden energy regions, the regions lying between the energy bands. To do this we consider electrons moving in the lattice according to wave mechanics as progressive electron waves whose wavelength λ (page 197) is

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2mE}} \tag{7-13}$$

if E is the energy of the electrons. Now it can be shown geometrically that only for certain definite wavelengths λ or certain energy values E the electron wave can penetrate the crystal. For different values of E or λ the electron wave is reflected by the lattice planes in such a way that it is extinguished by interference. The quantitative treatment of these considerations, which are only indicated here, leads naturally to the same results as the other two methods of approach. For this exact computation an expression for the three-dimensional periodic lattice potential is sought in which the electrons move. This expression is inserted in the Schrödinger equation (4-42), page 207. This equation has for solutions the eigenfunctions which describe the behavior of the electrons in the crystal, only for certain energy regions, i.e., for exactly the energy bands of our diagram Fig. 203.

In order to prevent a possible misunderstanding we want to emphasize one point. The broad energy bands of solids are pure electron energy bands and cannot be regarded as smeared bands of band systems of macromolecules. In molecules (page 364), electronic motion and atomic vibration are closely coupled with one another because an excitation of an electron produces in general a change in the molecular binding and as a result produces a vibrational excitation (Franck-Condon principle, page 370). Thus in a molecule an electron transition generally is associated with a change of vibration (rotation). A whole band system results from *one* electron transition because of the large number of vibrational and rotational energy changes coupled with it. If we disregard the internal vibrations in the molecular crystals, *crystal vibrations affect, in general, the whole crystal lattice. It seems evident that lattice vibrations are less influenced by the excitation of individual electrons.* Thus in a crystal the coupling between the motion of electrons and the vibration is very small. To an electron transition from one energy band to another in a crystal there corresponds a spectrum which might be regarded as a broadened spectral line and is usually called a spectral band. It should be kept in mind, however, that it has nothing to do with a molecular band or band system, because in the case of a spectral band of a crystal the characteristic property of a band spectrum, the participation of the vibration and rotation energy, is missing.

We return now to the energy bands of the crystal electrons. The distribution of the k levels (see page 438) over each energy band can be taken from the spectrum: its density is greatest in the center of the band and it decreases to zero at the edges of the band. The intensity distribution of a spectrum resulting from a transition from a sharp (inner) level to an energy band thus is similar to that of a broadened spectral line (the bell shaped curve, Fig. 103).

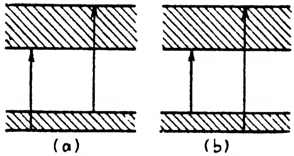


FIG. 205. Different cases of allowed transitions between two energy bands.

For the selection rule for optical transitions between energy bands it follows from the theory that transitions between all the different energy bands (different n's) are allowed, but with the important restriction that in optical transitions the quantum number k must not change. Since k is related to the momentum of the electron, this is essentially a consequence of the law of conservation of momentum. Surprisingly, the enumeration of k in the energy bands is not the same. If, as assumed in Fig. 205*a*, k increases from the lower to the upper edge of the energy band, then the short- and longwave optical transitions are shown by the arrows in Fig. 205*a*. The width of the corresponding spectral band is equal to the difference in the widths of the two combining energy bands. However, if k increases in one band from the lower to the upper edge of the band, and in the other band from the upper to the lower edge, then the possible short- and longwave transitions are shown by the arrows in Fig. 205*b*. In this case the width of the spectral band is given by the sum of the widths of the energy bands. In the case of metals, for example, both cases occur. We shall see on page 470 how important the k-selection rule is for understanding the luminescence of crystals.

A very nice experimental proof for the width and arrangement of the nonoccupied optical energy bands can be taken from X-ray absorption spectra. According to page 123, the longwave limit of the X-ray absorption continua originates from raising an electron to an unoccupied energy level near the ionization limit of the atom. The longwave structure of an X-ray absorption edge thus reflects directly the energy band structure in the crystal. This has been found to be in best agreement with the theoretical results developed above.

There is little to say about the optical spectra of solids associated with electron transitions so long as we speak of pure and ideal crystals. Absorption spectra usually result from transitions of electrons from the highest occupied energy band into one of the higher unoccupied bands. Depending on the different magnitudes of the interaction of the electrons, i.e. their exchange probability, we find broad continuous absorption bands in the ultraviolet and partly in the visible spectral regions for the metals, and very narrow bands, almost lines, for the crystals of the rare earths. This last result agrees well with our pictorial concept: the spectra of the rare earths result from transitions of electrons within the inner 4*f* shell which is shielded by the 5-quantum electrons (see Fig. 99) against external influences. These energy bands, consequently, are broadened only slightly. The large number of these transitions is due to the fact that the selection rules which are valid for isolated atoms are made invalid by the external influence of the interatomic electric fields in the crystal (see page 129).

Emission spectra of pure, ideal crystals can result from transitions of excited electrons from a normally unoccupied band to holes of a lower band, but these are very improbable, since the requirement of equal *k*-quantum numbers for the excited electron and the hole in the nearly filled lower band is almost never fulfilled. We shall become acquainted with the details of this on page 470).

7-9. Completely and Incompletely Occupied Energy Bands in Crystals. Electronic Conductors and Insulators According to the Energy Band Model

We can use the energy band model of the crystal electrons to draw conclusions about the electronic conductivity of crystals, conclusions which are as pictorially clear as they are important. Our aim is to give an atomistic explanation of the fundamental difference between metals and insulators. What are the conditions which determine whether a crystal is a good electronic conductor or an insulator?

In order that electric charge can be transported through a crystal by its electrons it is necessary that, as a result of the movement of electrons in the electric field to the positive pole, there be an excess of electrons on the positive side and a deficit on the negative side. Now if the highest band occupied by electrons is completely filled, and if we disregard the possibility of an electron jumping to a higher band (since this requires very high temperatures or optical excitation), then the accumulation of an excess of electrons on one side of the crystal is not possible. The migration of electrons toward one side of a crystal through a fully occupied band is possible only if simultaneously an equal number of electrons migrates toward the other side. One can understand that a unidirectional migration of electrons in a filled band is impossible, because the electrons would have to take energy from the field if they are accelerated by it. However, there are no energy states available for these electrons of higher energy in a filled band. Thus, *electron conduction is not possible in crystals with completely filled energy bands. These crystals therefore are called electronic insulators.*

On the other hand, *if in a crystal the highest energy band which contains electrons is* not *completely filled, then, according to the above statements, unidirectional migration of electrons is possible. Consequently, metallic conductors are characterized by the property that the highest of the occupied energy bands is incompletely filled.*

How does this theoretical result agree with our empirical knowledge of the highest occupied energy bands of the metals? According to our discussion on page 439, each energy band of a crystal consisting of N atoms has "room" for 2 N electrons, i.e., for two electrons per atom. According to the Pauli principle (page 169), each k state can be occupied by two electrons with oppositely directed spin. We also know that the highest energy band of a crystal is occupied by the valence electrons of the constituent atoms. Consequently, it is clear that the monovalent metals are electrical conductors, because they have only one electron per atom in the highest energy band, which thus is only half filled. However, according to this simplest concept, bivalent metals should be insulators, in contrast to experience. Now in metals the interaction of the electrons which determines the width of their energy bands (and which, according to page 432, also causes the metallic binding of the atoms in the crystal) often is so large that the upper energy bands of the metals partially overlap. Since the state to which the system tends is always that of the smallest potential energy, the $2N$ valence electrons will distribute themselves in the manner shown in Fig. 206. In this figure the bands are shown one beside the other. Instead of completely filling one band and leaving the other unfilled, the $2N$ valence electrons thus fill both bands up to the same level. *As a result of this overlapping of the higher bands in a crystal with large interaction between the electrons, the highest occupied bands of the bivalent metals are not completely filled and these solids thus show the electronic conductivity which is characteristic for metals.*

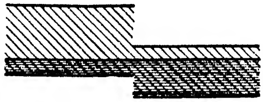


FIG. 206. Overlapping of the highest occupied energy bands of a bivalent metal and occupation of the two bands by electrons. For clarity the two bands are shown side by side.

That the situation presented here is not only a hypothesis but that it corresponds exactly to reality has been shown by a very beautiful experiment of Skinner. The experiment concerns the X-ray emission spectra of various metals. We know from page 438 that the lower electronic states in a crystal can be regarded as sharp levels because of the small disturbance of the innermost electrons of the atoms by the lattice. The X-ray lines of lighter elements which are emitted after ionizing an electron of the K shell (page 117), correspond to transitions from the upper broad energy band (so-called conduction band) to the K level which is very sharp in comparison. The width and intensity distribution of the emitted X-ray band thus gives us a direct picture of the width of the upper energy band and its electron population. Densitometer curves of these X-ray emission bands are shown in Fig. 207 for the monovalent lithium metal and the bivalent magnesium metal. *Only half of a bell-shaped curve is clearly visible in the first case, and this is pictorial evidence that the upper band of lithium is only half-filled with electrons. In the case of the bivalent magnesium, on the other hand, it can be easily recognized that the X-ray emission band originates from the transitions from* two *partially occupied, overlapping energy bands to the same lower state.* This is in complete agreement with our discussion given above. In a similar way the energy bands of the alkali-halide crystals and those of a number of oxide crystals have also been studied and attributed to definite electron transitions.

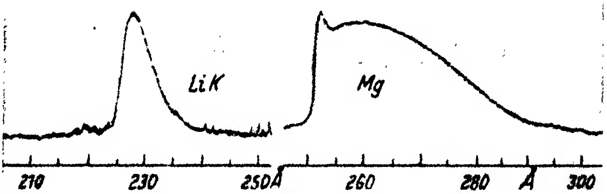


FIG. 207. Densitometer recordings of the X-ray emission bands of a monovalent (Li) and a bivalent metal (Mg), as evidence for the correctness of the band arrangement in Mg according to Fig. 206. (*After Skinner.*)

We conclude with a remark about the temperature dependence of the electronic conductivity. In the ideal nonvibrating crystal, the coupling between the electrons in the conduction band and the lattice built up of their ions is very small and the conductivity is correspondingly large. The lattice vibrations tend to reduce the electron mobility because the electrons collide more often with ions and are scattered irregularly. As the amplitude of the lattice vibrations increases with increasing temperature, the migration of the electrons through the crystal is inhibited to an increasing extent. The electric resistance of the metals increases with increasing temperature, because the electrons must "work their way" through a potential field which varies with respect to space and time. The empirical law according to which the electronic conductivity at low temperatures is proportional to T^{-5} thus appears to be explainable in a qualitative way. On the other hand, a conduction phenomenon which cannot be satisfactorily explained from the theoretical point of view is that of supraconductivity. This is the remarkable effect that a large number of metals and metallic compounds seem to lose their electrical resistance completely in the vicinity of the absolute zero of temperature. According to Heisenberg, this phenomenon seems to be associated with a lattice-like arrangement (produced by electrostatic repulsive forces) of those few high-energy conduction electrons which are not yet "frozen" in their lowest possible energy states.

7-10. The Potential Well Model of a Metal. Work Function, Thermionic Electron Emission, Field Emission, Contact Potential

In our discussion of solids and particularly of metals from the atomistic viewpoint we have neglected surface effects. We shall now interest ourselves in those phenomena, and especially in the passage of electrons through the surface layer of a crystal into free space (vacuum or air) or into another metal.

In this respect we must *think of the electrons in the conduction band as bound not to the individual lattice ions but to the crystal as a whole and that, consequently, for an electron to escape out of the crystal, energy must be expended against this binding force.* The energy required to release an electron from a metal into free space is called the work function Φ.

Thus the potential energy of an electron is greater in free space by the amount Φ than it is for the electron in the crystal. We can represent this behavior by a potential curve in the customary manner and thus obtain the potential well model of a metal (Fig. 208). In this representation we are concerned only with the quasi-free conduction electrons. Only their energy levels are, therefore, indicated in Fig. 208. We see that the potential of an electron at rest in the interior of the metal (lower edge of the conduction band) is smaller by an amount Φ_0 than it is outside the metal. However, this potential difference is not the effective work function for the release of an electron from the metal. The effective work function Φ is much smaller than Φ_0 because the energy levels within the potential well are filled by electrons up to a height δ. According to Fermi statistics (page 238) electrons in the highest state already have a zero-point energy equal to δ. The effective work function for those electrons which are most easily released from the metal therefore is

$$\Phi = \Phi_0 - \delta \tag{7-14}$$

The behavior described here holds exactly for absolute zero, whereas at higher temperatures there are always a certain number of electrons in higher, formerly unfilled states.

The potential well model gives us a very pictorial explanation for the emission of electrons by an incandescent metal. It is clear that electrons with kinetic energies $E_k > \Phi$ can leave a metal spontaneously; in the classical picture they evaporate. In 1908, Richardson deduced from measurements a formula for the temperature dependence of the saturation current i which is obtained when the released electrons are drawn off by a sufficiently high electric field. His result is the famous Richardson equation,

$$i = AT^2e^{-\Phi/kT} \tag{7-15}$$

A is a constant of the metal which is theoretically equal to 120 if the current density is computed in amperes per square centimeter of the incandescent surface. However, for the large number of materials investigated, the value of A varies between 10^{-3} and 10^4, if the value of the normal work function is used for Φ. This discrepancy is explained by the fact that actually the work function Φ of different materials changes in different ways with the temperature. However, customarily the measurements are represented by the expression (7-15) with constant Φ, but different A. The Richardson equation has been derived by Schottky, Dushman, and others from the classical theory, before the advent of the quantum theory, on the basis of a normal evaporation of the electrons from the heated metal.

FIG. 208. Potential-well model of a metal. Φ = effective work function, ϕ_0 = work function for electrons in the lowest energy state of the conduction band, δ = Fermi surface = zero-point energy of the electrons in the highest occupied electron levels.

Compared to the exponential term $e^{-\Phi/kT}$, where Φ is the effective work function, the term T^2 is rather unimportant. It seems clear that the electron current has to be proportional to the term $e^{-\Phi/kT}$, because this quantity gives the fraction of the electrons which have at the temperature T the energy Φ. The quantum-mechanical derivation from the electron theory of metals leads to exactly the same equation.

Since the work function Φ occurs in the exponent of Richardson's equation, a slight reduction in the work function causes a large increase in the number of emitted electrons. The effective work functions for several technologically important metals are listed in Table 16.

Table 16. The Work Function of Several Metals

Metal.............	Cs	Ba	Th	W	Pt
Φ electron (volts)...	1.8	2.1	3.35	4.5	6.2

We see that the variation in the work function corresponds to that of the ionization potential of the atoms forming the metal. Cesium, for example, has the smallest work function and likewise the smallest ionization potential of all elements (see Table 2, page 30). This relation might be expected from the quantum theory, since *the work function* (actually not Φ, but Φ_0) *can be regarded as the ionization energy of the atom when influenced by the crystal lattice.*

The strong dependence of the thermionic emission on the work function is used widely in technical applications. Since the cathode cannot be made of pure cesium, barium, or thorium for various technological reasons, tungsten, which can withstand high temperatures, is used. However, its work function is lowered by depositing on it a layer of cesium or thorium, monatomic, if possible, which results in a large increase of the thermionic electron emission. In the first method, the tungsten cathode is operated in a rarefied cesium vapor atmosphere, by which a monatomic cesium layer is formed on the tungsten. To produce the thorium coating, in the manufacturing process the tungsten is mixed with thorium oxide and this is reduced by heating so that the thorium diffuses toward the surface, thus producing a monatomic thorium layer on the tungsten. This *thoriated tungsten cathode* has the great advantage that it can be used in high vacuum. Surprisingly, by properly coating the tungsten with Cs or Th atoms, a work function is obtained which is only slightly higher than 0.7 or 2.6 ev, respectively, and thus is noticeably *less* than that of pure metallic cesium or thorium. This can be explained by the atomic theory and is confirmed by experiment as being due to the fact that the Cs and Th atoms adsorbed on the tungsten surface are polarized with the positive pole of the resulting dipole directed toward the tungsten surface. In this way the positive charge is stronger and the negative electron cloud less tightly bound than for symmetrical binding. The separation of an electron from this type of adsorbed atoms therefore requires an amount of energy smaller than the normal effective work function. This phenomenon is an excellent example of the atomistic explanation of even complicated details of a technically important effect.

Another application of the model of the potential well is the field emission of electrons from cold metals. It is the phenomenon that at electric field strengths above some million volts per centimeter electrons are emitted from cold metals. Such high fields occur particularly near sharp points of electrodes. The effect of field emission, therefore, seems to play an important role for the initiation of certain types of spark discharges.

If we discuss field emission by means of the potential well model, we have to superimpose on the potential of the binding forces the potential of the strong external electric field. The result is the distorted, asymmetrical shape of the potential well edges which is indicated in Fig. 209. Such a potential distribution implies that the potential of an electron on the positive side of the field outside the metal would be smaller than inside the metal. Therefore, electrons can penetrate the potential wall (tunnel effect) at the place indicated by the dotted line and escape into free space. It should be noticed that this explanation of field emission corresponds closely to the explanation of the preionization of atoms (page 186) by a distortion of the potential curve and tunnel effect. Since, according to page 236, the penetration probability increases with decreasing height and width of the potential wall to be penetrated, one can draw three interesting conclusions. First, the field emission should increase greatly with increasing field strength because, as a result of distorting the edge, the width of the potential wall is reduced. Second, with increasing temperatures the number of escaping electrons should increase because the electrons in the higher states (bands) in the potential well have lower and thinner potential walls to penetrate (transition to thermionic emission). Third, similar to the case of thermionic emission, the field emission should increase strongly, or begin at lower field strengths, if the effective work function (i.e., the height of the potential wall to be penetrated) is reduced by suitable adsorbed layers on the cathode surface.

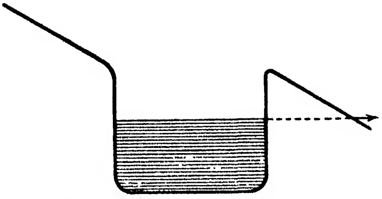


FIG. 209. Distortion of the potential wall in a strong electric field. This makes possible the field emission of electrons from the metal by way of the tunnel effect (dotted arrow).

The increase of field emission with increasing temperature does not seem to be of particular interest and has not been measured. The increase with increasing field strength has been measured and found to be in agreement with the expectation: the field emission increases exponentially with the field strength. Reliable quantitative measurements are difficult, because adsorbed surface layers and microscopic irregularities cause deviations from the expected behavior. For instance, for thoriated tungsten, an appreciable field emission has been found for fields of 1 million volts/cm, though theoretically field emission was not expected below 10^7 volts/cm. However, careful removal of all surface layers and consequent outgassing of the electrode surface led to a decrease of the field emission. The remaining discrepancy seems to be due to the fact that the actual field strength near microscopic peaks of the surface is larger than the computed one. Conclusive proof for this has been recently found by electron-microscope measurements of emitting metal points. The field emission of cold metals here actually was found to begin at 3×10^7 volts/cm.

As a last example for the application of the potential well model we discuss the contact potential between two different metals. Two metals may be brought into contact and then slightly separated. Between them, we can then measure a potential difference, the contact voltage. Each of the metals may be characterized by a potential well with different values of Φ_0, Φ, and δ (Fig. 210). If the two metals are in contact, they have the same potential. This means that electrons would flow from one to the other, until both electron surfaces have the same height, as it is indicated in Fig. 210. However, now an electron at the surface of metal II has a higher potential energy than one at the surface of metal I, because the work function of metal II in our example is larger than that of metal I. The difference of the potential energies of these electrons is just the potential difference between the two surfaces, i.e., the contact voltage. Its value is evidently equal to the difference of the effective work functions of the two metals. This contact voltage exists also between the two separated metals. Because of the difference of the work functions, more electrons will leave the metal I than II, until the two metals have a potential difference which prevents further electron current. Again we arrive at the result that the contact voltage is Φ_{I} minus Φ_{II}. This theoretical result is in agreement with experiment, provided that the measurements are made with extremely clean and outgassed metal surfaces. In most cases, surface layers of various kinds cause complications. It is obvious that, conversely, the effect of such surface layers, loose contacts, etc., can be studied by means of analogous considerations with potential well models. This has been done successfully and is of interest for electrical engineering.

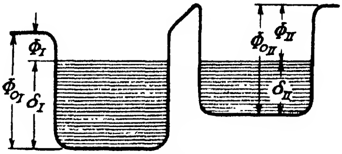


FIG. 210. Explanation of the contact potential between two different metals I and II.

7-11. Ferromagnetism as a Crystal Property

Ferromagnetism is one of the crystal properties which has been explained by quantum mechanics. That it is actually *a property of the crystal and not of the single atoms* follows from the fact that, for example, iron vapor or iron compounds do not exhibit ferromagnetism (see page 152), whereas the Heusler alloys formed from the nonferrous metals copper, manganese, and aluminum are ferromagnetic. Moreover, certain unmagnetic crystals can be made ferromagnetic just by changing their lattice structures.

The most remarkable property of ferromagnetic materials is that even a very small external magnetic field produces a large magnetization. Moreover, the magnetization does not increase steadily by increasing the field. A large-scale picture of the hysteresis curve (Fig. 211) shows that the magnetization increases in discontinuous steps, the so-called Barkhausen steps. This stepwise, large increase of the magnetization under the influence of a comparatively weak external magnetic field can be explained only by the assumption that *even without an external field entire crystal domains are magnetized (spontaneous magnetization) and that the magnetic field only serves to overcome the resistance which prevents the instantaneous orientation of the magnetic domains in the direction of the external field.*

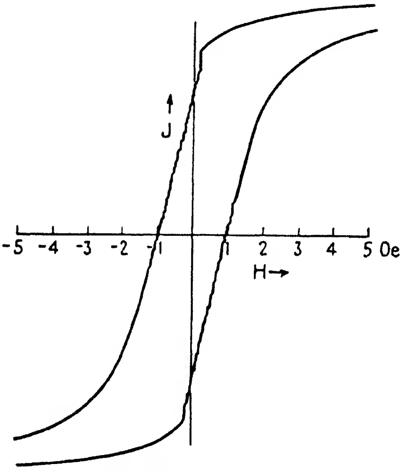


Fig. 211. Magnetization curve with Barkhausen steps. (*After Becker and Döring.*)

How can we explain this strong spontaneous magnetization of whole regions in a ferromagnetic substance even without an external field? It can only result from the parallel orientation of the magnetic moments of all or almost all conduction electrons of the highest occupied band in the individual domains in the metal under consideration. Normally, according to page 439, each state, characterized by the quantum number k, of the highest energy band is occupied by 2 electrons. According to the Pauli principle (page 169) these two electrons must have opposite spins so that their magnetic moments compensate each other. Now if all magnetic moments are oriented in the same sense, then each k state can be occupied by only one electron. Then for the same number of electrons twice as many k states are required as in the nonferromagnetic case. This number of k states is available in metals, because their highest energy bands are generally only half-filled (page 443). However, the electrons must, on the average, now occupy higher k states of the conduction band than normally, so that their average zero-point energy is higher too. Now such a spin orientation can occur spontaneously only if energy is *gained* in the process. This means that the energy gain due to the parallel orientation of the spin momenta must overcompensate the increase in the average zero-point energy of the conduction electrons.

The answer to the problem of how a parallel orientation of the spin momenta can occur and lead to a lower potential energy of the system than the normal state of antiparallel spin momenta, was given by Heisenberg in a famous application of quantum mechanics. According to page 231, the energy of interaction of two electrons 1 and 2 which may belong to the ions *a* and *b*, consists of an electrostatic part which is independent of the spin orientation, and the exchange energy which depends on the mutual spin orientation. It is positive or negative depending on whether the exchange integral

$$A = \int \psi_a(1)\psi_b(2)\psi_a(2)\psi_b(1)\left(\frac{1}{r_{ab}} - \frac{1}{r_{a2}} - \frac{1}{r_{b1}} + \frac{1}{r_{1,2}}\right) d\tau \qquad (7\text{-}16)$$

is positive or negative. Thus, *if this exchange integral of the eigenfunctions of electrons with spin in the same direction is positive, then this ferromagnetic state of the crystal is energetically more favorable than the nonmagnetic state with compensating magnetic moments.* From (7-16) it can be seen that A will be positive if the average distance between the two interacting electrons, $r_{1,2}$, is small and the distances between the nuclei and electrons r_{a2} and r_{b1} are large. Accordingly we expect, after Slater, a spontaneous orientation of the spin moments of the quasi-free electrons in a crystal, if two conditions are fulfilled:

1. The atoms forming the crystal must not have closed shells, so that the spin momenta of the external electrons are not mutually compensated. Furthermore, these electrons should occupy a state of high orbital quantum number in which the electrons are relatively far from their nuclei.
2. The radius of the electron shell must be small compared to the lattice distance, since only then the integral over the product of the electron eigenfunctions divided by $r_{1,2}$ has a large value.

Condition 1 is fulfilled for all transition metals in the periodic table, whereas condition 2 holds only for the rare earths and the group, iron, cobalt, and nickel. Though the exchange integral is positive for the rare earths, its value is so small that even a slight thermal agitation is able to counteract the spontaneous orientation of the spin magnets in the crystal. The crystals of the rare-earth elements thus are ferromagnetic only in the neighborhood of absolute zero. Their Curie points, above which their ferromagnetism disappears, is far below room temperature. Only for the metals iron, cobalt, and nickel is the exchange integral large enough to cause a very stable spontaneous orientation of the spin momenta. Only these metals therefore are ferromagnetic up to their fairly high Curie points of 360 to 1000°C. Figure 212 shows the value of the exchange integral A (7-16) with its dependence on the ratio of the interatomic distance to the radius of the unclosed electron shell. This explanation of ferromagnetism by the atomic theory is confirmed by the fact that ferromagnetism can be produced by increasing the lattice distance of the nonmagnetic metal manganese. According to Fig. 212, its exchange integral is only slightly negative and becomes positive if we increase the interatomic distance by incorporating nitrogen, i.e., by converting the metal to manganese nitride.

A detailed discussion of the behavior of the various modifications (lattice structures) of iron by Dehlinger revealed that apparently not only the exchange energy between an atom and its immediate neighbors is of importance. Because the outer electrons of metals are not well localized (see page 432) or, in other words, because the eigenfunctions are extended over the volume of many atoms, the interaction between a particular atom and its more distant neighbors has to be taken into account and may determine whether, in a particular case, parallel spins are to be expected or not. Consequently, a parallel orientation of the spin momenta of the conduction electrons, and thus ferromagnetism, is to be expected if the exchange integral obtained by summing the exchange integrals between an arbitrary atom and all other atoms of the lattice is positive. With this refinement not only the magnetic behavior of the various modifications of iron can be understood, but also the complicated behavior of the many magnetic alloys which have been discovered in the past 20 years.

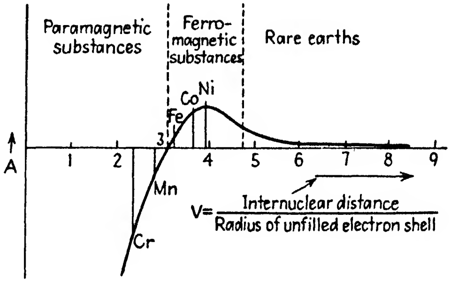


FIG. 212. Curve giving the value of the exchange integral (which determines ferromagnetism and Curie point) for a number of ferromagnetic or nearly ferromagnetic metals as a function of the ratio of interatomic distance to the radius of the incomplete electron shell. (*After Becker and Döring.*)

The spontaneous magnetization (without an external field) of certain crystals can thus be understood on the basis of the atomic theory. That a piece of iron in general is not magnetic without an external field, is due to the fact that this spontaneous magnetization occurs only in small crystal domains of 100 to 10,000 atoms diameter, and these elementary domains are usually oriented completely at random in the crystal. An external magnetic field is thus required in order to orientate the domains. That such an orientation is possible means that these elementary domains can be rotated in the solid crystal though they offer some resistance against an arbitrary rotation. This peculiar crystal property has been studied carefully by measuring the hysteresis curves of single crystals of iron. For these studies the external magnetic field was directed along an edge, along a surface diagonal or the volume diagonal of the cubic iron crystal. Figure 213 shows the results. If the magnetic field acted along a cube edge (100), the magnetization increased rapidly with increasing orienting field to the saturation value M_s. Evidently only a very small orienting force is sufficient to turn the spontaneously magnetized domains in the direction of the cube edge. If the direction of the field coincided with a surface diagonal (110), the magnetization increased rapidly up to the value $M_s/\sqrt{2}$, and with the field in the direction of the cube diagonal (111), up to the value $M_s/\sqrt{3}$. It then rose slowly to the saturation value M_s. This behavior is understandable if it is assumed that all spontaneously magnetized domains do not offer any appreciable resistance against being rotated in the direction of the cube edge closest to the field direction, but that it requires more force to turn the spins from this preferred direction into the actual field direction. After the external field has been removed, the domains will reorient themselves in the direction of the nearest cube edge and will remain there if the thermal agitation is not too great. This is the explanation of the well-known remanent magnetism.

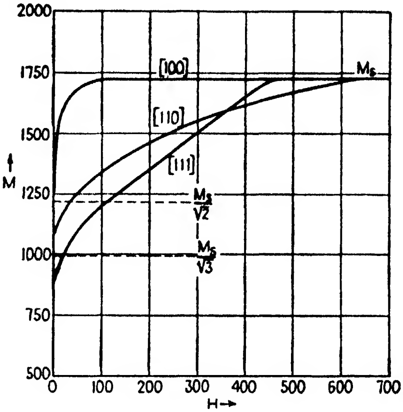


FIG. 213. Magnetization curves of an iron single crystal for orientation of the field in different crystal directions. (*After Honda and Kaya.*)

For the usual polycrystalline iron, and particularly for iron as it is treated in many manufacturing processes, the behavior is much more complicated than for single crystals. This is chiefly because the inner stresses produced in working the iron give rise to preferred directions of orientation for the domains. However, even here the complicated phenomena of ferromagnetism can be understood in principle from the well-known rules of atomic physics.

Recently the interesting question has been studied how large a Fe microcrystal must be in order to exhibit ferromagnetism (in contrast to the Fe atom). According to the experiments of König the microcrystal must have at least 64 elementary cells (page 425). If the surface atoms which are not essential for the spin orientation are disregarded, this number reduces to 8, i.e., to a body which also seems to play a special role as a crystal seed. Here we have apparently arrived at the borderline between atomic complexes and solid bodies, i.e., microcrystals.

One more important property of metals should be mentioned at this point, though it cannot yet be explained by atomic physics: *supraconductivity* (page 445). This strange effect also depends on a special electron orientation and apparently, just as ferromagnetism, is possible only in certain crystal lattices. It seems that, because of a special "configuration," the conduction electrons can move through the supraconducting metallic lattice practically without any resistance. Since there is no satisfactory atomic explanation for this important and highly interesting phenomenon, we have not devoted a special section to it, but only mention it in concluding the discussion of ferromagnetism which also depends on a special electron orientation in certain crystal lattices.

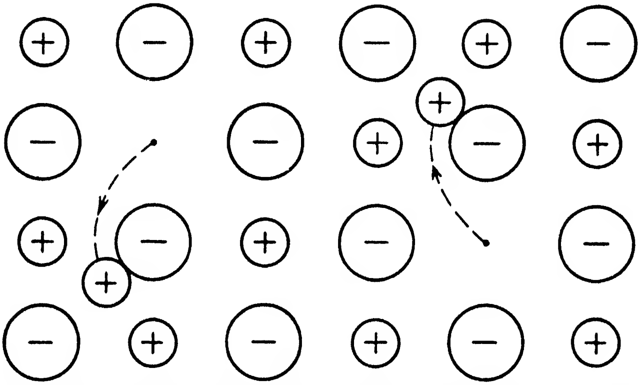


FIG. 214. Positive interstitial ions and vacant lattice places in an ionic crystal: Frenkel defects.

7-12. Lattice Defects and Their Bearing on Diffusion and Ionic Conduction in Crystals

Ferromagnetism and supraconductivity seem to be structure-sensitive properties (see page 422) in so far as a particular lattice is necessary for these phenomena. We shall now discuss those properties of solids which, like diffusion and ionic conductivity, depend essentially on lattice defects. In an ideal crystal each constituent is in its own potential trough and any migration of a particle—for example, the exchange in the location of two particles (atoms, ions, or molecules according to the lattice type, page 423) in the case of spontaneous diffusion—requires energy in order to overcome the potential wall. Spontaneous diffusion therefore is strongly temperature dependent. Because it increases exponentially with the temperature, it is difficult to observe at room temperature.

According to page 422, each real crystal has a large number of lattice defects, such as vacant lattice places which would normally be occupied by a constituent, and the so-called interstitial ions. During the growth of the crystal, according to Fig. 214, a normal lattice ion can "erroneously" be "frozen" in an interstitial place, i.e., it did not get into its right potential trough but, by deformation of the crystal, has made a place for itself in the neighborhood of the correct place. By this process there is produced in the crystal an empty place and an interstitial ion. This type of defect is called a *Frenkel defect.* Also during crystal growth or by subsequent migration, ions of both signs (since the crystal is electrically neutral) may have come to the crystal surface; empty places then remain in the crystal (Fig. 215). These defects are called *Schottky defects.* Both types of defects enable lattice constituents to have a much greater mobility, since *with lattice holes the diffusion can progress stepwise through the lattice holes.* This is called *hole diffusion,* and the process is called *defect conduction.* Obviously much less activa-

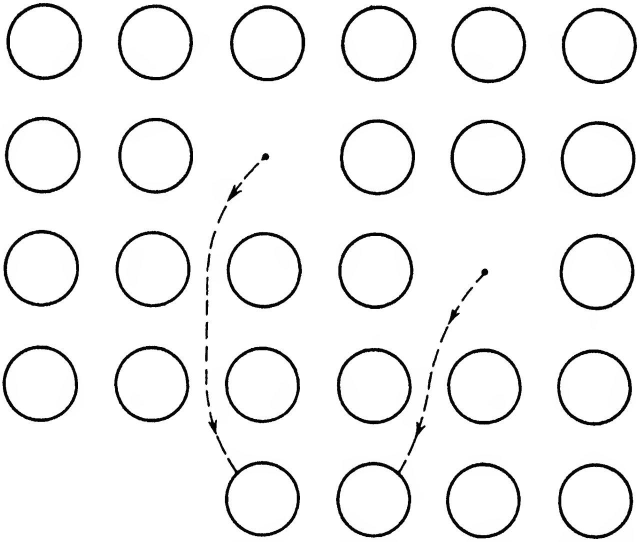


FIG. 215. Positive and negative vacant lattice places: Schottky defects.

tion energy is required for hole migration or the motion of interstitial ions than for an exchange of normal ions in a lattice. Since the number of lattice defects depends on the incidentally more or less regular crystal growth and its later treatment (heating, mechanical stress, etc.), the defect diffusion varies greatly between different samples of the same type of crystal. This behavior is in contrast to the characteristic high temperature diffusion mentioned above, which is independent of lattice defects.

The study of the diffusion of atoms or ions in their own crystal, for instance of lead atoms in lead metal, has become possible, as mentioned on page 309, by the radioactive tracer method (diffusion of radioactive atoms or ions). The diffusion of two solid materials into each other, when in contact, e.g., that of silver into gold and conversely, is one of the migration phenomena which is easier to study. At low temperatures the diffusion in solids is exclusively due to vacant lattice places and interstitial atoms. However, it is obvious that this migration process is also strongly temperature-dependent. With increasing temperature higher and higher lattice vibrations are excited and lead to a "loosening up" of the lattice which facilitates the migration of lattice constituents. It is, finally, understandable that the velocity of diffusion of one solid into another is the larger, the less the difference in the diameters of the two kinds of atoms. An occasionally very high concentration of lattice defects plays an essential role in the semiconductors and phosphors which we shall discuss in detail later. In these crystals, either one of the lattice constituents is present in excess (e.g., Zn in ZnO), or foreign atoms are introduced into the lattice (so-called impurity-activated crystals). In both cases we find interstitial atoms as well as lattice vacancies.

It is evident that gas atoms or ions can also diffuse into solids. The diffusion of hydrogen through palladium and iron, to cite only two prominent examples, is well known to gas discharge physicists.

The great importance of diffusion processes for the practical use of metals becomes apparent from the example of rusting of iron. As soon as metallic iron is covered by a layer of rust, it is no longer possible for oxygen or OH radicals to attack the metal since this is now protected by a layer of another solid, rust. Further rusting is possible only by diffusion of either oxygen or metal through the crystalline layer of rust. Because of the smaller radii of the positive ions it seems that the diffusion of the iron ions to the surface of the rust layer is more important than the reverse diffusion of oxygen. The number of similar examples can be increased arbitrarily.

We have not taken into consideration up to now that the lattice constituents of an ionic lattice are ions of opposite polarity. Thus, if we apply an electric field to an ionic crystal, its ions migrate to the electrodes of opposite sign, so that ionic crystals exhibit an electric conductivity which is ionic just as is the conductivity in electrolytic solutions. It is obvious that this ionic conductivity has nothing to do with the migration of electrons and the corresponding electronic conductivity in certain crystals (semiconductors) which we shall discuss later. Because the ionic conductivity depends on the migration of lattice ions, we have the empirically well known negative temperature coefficient of resistance of ionic crystals: the resistance decreases with increasing temperature, because the lattice loosens up. *At low temperatures the ionic conductivity is strongly structure-dependent, and the migration occurs almost exclusively via the vacancies and interstitial lattice places. At high temperatures, on the other hand, the structure dependency vanishes and, because of the loosening of the lattice, the ions migrate by exchanging places in the lattice.*

That the ionic conductivity increases exponentially with the temperature is to be expected for an exchange of places by overcoming potential walls. For a particular sample of a NaCl crystal, according to Smekal, the conductivity could be represented by the formula

$$\sigma = 0.1e^{-10,000/T} + 10^6e^{-24,000/T} \qquad (7\text{-}17)$$

The difference in the order of magnitude of the constants, 0.1 and 10^6, shows that only the 10^7 part of the lattice constituents participates in the conductivity represented by the first term. This term consequently describes the conductivity due to lattice defects. The second term represents the conductivity resulting from the migration of normal lattice ions. In agreement with this explanation the activation energy of the defect conduction is only 40 per cent of that of the second term (10,000 to 24,000). The average height of the potential walls for both processes is 0.8 and 2.0 ev, respectively. Because of the small height of the potential wall (activation energy) in the exponent, the first term of the conductivity formula predominates at low temperatures in spite of the small numerical factor, whereas at high temperatures the second term, because of its large factor, predominates. Since, in the general case, holes, interstitial ions, and normal ions of both signs contribute to the conductivity, the temperature dependence of the conductivity of ionic crystals should be represented by a formula with four terms. Practically, however, it can be represented by two terms. Because of the different activation energies of the various partners and the limited temperature regions in which it is feasible to make measurements, only two terms play an important role, unless the accuracy of the measurements is increased considerably.

7-13. Light Absorption and Electron Motion in Halide Crystals. The Primary Photographic Process

In the last section we were concerned with the movement of lattice constituents as well as of foreign atoms or ions in the crystal lattice and learned of the decisive influence of lattice defects on this migration. We shall now discuss the effect of light on ionic crystals, and here we shall deal for the first time with the *migration of electrons in nonmetallic crystals*. The fundamental investigations on the effect of light on ionic crystals were carried out by Pohl, Hilsch, and coworkers on "colored" alkali-halide crystals. In addition to their general interest, they are important for understanding the fundamental processes of photography.

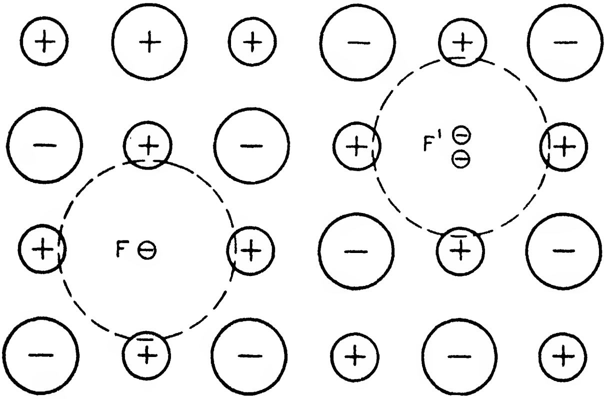


FIG. 216. Schematic representation of the probable nature of F and F' centers in colored alkali-halide crystals: one or two electrons in the lattice place of a missing negative ion.

If transparent alkali-halide crystals are irradiated with sufficiently short ultraviolet light, with X- or cathode rays, or if they are heated in the vapors of their characteristic metals (e.g., NaCl in Na vapor), the crystals become colored. The characteristic color for each crystal varies, for instance, from yellow for LiCl to blue for CsCl. The absorption band corresponding to this color, which is missing in the untreated crystal, has been called the F band (Farb band or color band) by Pohl. The corresponding absorbing centers are called F centers. It required many years of intensive studies to explain the nature of these F centers, and it is not possible to present here all the experiments which support the presently accepted explanation. According to our present knowledge, an F center is not, as was formerly believed, simply a metal atom produced by attaching a free electron to an arbitrary metal ion of the lattice. *An F center seems to consist of an electron which has taken the place of a missing negative halogen ion (Fig. 216) and then, as a non-*

localized electron (page 415), alternatingly forms an absorbing Na *atom with each of the six surrounding* Na *ions.* Its excitation energy of about 3 volts, as computed by wave mechanics, agrees well with the wavelength of its absorption band. To ionize an *F* center, i.e., to release the electron from its potential well, an energy about 0.1 volt higher than its excitation energy is required. The following experiment is in agreement with this result: If a colored crystal at a sufficiently low temperature is irradiated with the wavelength of its absorption band, then the *F* electron excited by the absorption remains bound to its place and the crystal thus remains an insulator. On the other hand, at elevated temperatures the energy lacking for the complete release of the *F* electron is supplied by the thermal vibrations of the lattice and the *F* electron is raised to the previously unoccupied conduction band. Here, as a free electron, it is responsible for electronic conductivity if an electric field is applied to the crystal (Fig. 217); the crystal exhibits "photoelectric conductivity."

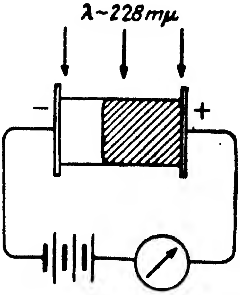


Fig. 217. Diagram of the Pohl experiment for the electric measurement of the electron migration in a colored KH–KBr mixed crystal. By ultraviolet irradiation electrons are liberated and migrate in the electric field from one vacant negative hole to the other toward the anode, thus establishing an electric convection current.

Pohl furthermore found that the destruction of *F* centers by irradiating in the *F* band at not too low temperatures produced not only the expected reduction of the absorption intensity of the *F* band, but also the appearance of a new longer wave absorption band which he called *F′* band. The nature of these *F′* centers which are responsible for this longwave absorption has been explained in the following way: *the F′ center is supposed to be an F center that has captured an additional electron (Fig. 216), or, in other words, two instead of one electron in the place of one missing* Cl^- *ion.* The absorption of the *F′* bands lies at longer wavelengths than that of the *F* bands, because it requires only a small amount of energy to separate the loosely bound second electron. The formation of *F′* centers according to the present explanation seems clear: an *F* electron, released from its potential trough by the absorption of light with the cooperation of the thermal lattice energy, is captured by another *F* center which is then transformed into an *F′* center. In agreement with this explanation, two *F* centers are destroyed for each quantum absorbed in an F band (quantum yield 2). One is destroyed directly by absorption and one is transformed into an *F′* center. The fact that the mean free paths of the photoelectrons released from the *F* centers is inversely proportional to the density of *F* centers is further direct confirmation of the explanation of the F centers as presented here. Furthermore, in general it is not possible to produce more than 10^{18} F centers per cubic centimeter in a crystal. This is further evidence in favor of the present theory, because in general it is not possible to form more F centers than there are Cl^- vacancies. Consequently, *the production of F centers consists in the introduction of free electrons into the halogen vacancies in the lattice.* The free electrons are produced by the absorption of ultraviolet light or X-rays or by collisions of cathode-ray electrons. If the coloration is produced by heating the crystal in metal vapor, the metal ions attach themselves to the crystal surface (together with halogen ions which migrate outward and thus produce vacant lattice places), while only their electrons migrate into the crystal. That actually an electron migration into the crystal produces the F centers, and thus the coloration, was shown by Pohl in his famous electrical experiment. If a potential difference of several hundred volts is put across a completely transparent KBr crystal and the lattice "loosened" by raising the temperature to 400 to 600°C, then electrons migrate from the negative electrode into the crystal and fall into Br^- vacancies, thus producing absorbing F centers. A blue cloud can be seen moving into the crystal, which can be moved back and forth as the polarity of the electric field is reversed (Fig. 218). Pohl called this production of F centers by introducing electrons into the crystal "*additive coloration*" in contrast to the coloration of crystals by irradiating them with X-rays. In this last case the electrons can return to their original places when the crystal is heated or irradiated with visible light. *The crystal can be bleached, in contrast to an additively colored crystal.* All these experiments with alkali-halide crystals have not only revealed a number of interesting processes in solids but they have also clearly shown *that lattice defects are of decisive importance for the effects of light on ionic crystals and the behavior of electrons in such crystals.*

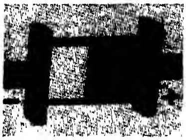


FIG. 218. Pohl's experiment by which the migration of electrons in an alkali-halide crystal is made visible. At 600°C the electrons in a KBr crystal migrate so quickly toward the positive pole that the rear boundary of the "blue cloud" formed by the F centers moves visibly toward the anode. (*After Pohl.*)

The same situation holds for the basic atomic processes of photography which are so complicated because of this structure dependency. Free electrons are primarily produced by absorption of light in the AgBr microcrystals of the photographic emulsion just as in the alkali-halide crystals. After filling the lattice vacancies, these electrons attach themselves to Ag^+ ions and form absorbing Ag atoms. The decisive but not yet entirely understood process is now that the individual Ag atoms coagulate and form colloidal-like silver nuclei. After the development of the "latent" image, i.e., after the chemical reduction of further Ag atoms, these silver nuclei cause a blackening in the plate at those places which had been irradiated.

Theoretically, in producing the latent image, each absorbed photon can release one electron which attaches itself to an ion and forms a silver atom. This quantum yield of 1 can actually be obtained in the elementary photographic process according to measurements of Eggert and Noddack. In general, the quantum yield depends on the illumination during the exposure since too strong exposures produce a reformation of AgBr which reduces the effective quantum yield.

That the addition of easily ionizable molecules, such as Ag_2S, increases the sensitivity of the photographic emulsion can be understood without further discussion. It is also clear that a certain minimum energy is necessary to separate the electron from the halogen ion. This energy determines the wavelength limit beyond which the emulsion is no longer sensitive. Photographic emulsions are always sensitive to shortwave light but, in general, they are not sensitive to longwave red light. It is known, however, that photographic emulsions can be sensitized for longwave light by adding certain dyes which are able to absorb longwave light.

Although one has been able to explain to a certain extent the elementary photographic process, one cannot yet fully explain the development of the latent image. It is known that the number of primary Ag atoms which form the latent image is increased by a factor up to 10^8 by chemical reduction of silver atoms, and that the existence of silver nuclei is necessary for this chemical development. Yet nothing definite is known about their structure or effects. In particular, it is not yet known whether the nucleus consists of colloidal silver or, as many investigators believe, regularly built micelles in which foreign atoms play an important role. The atomic physicist is still looking forward to more complete empirical results.

7-14. Electronic Conduction in Semiconductors and Related Phenomena. Barrier Layers and Rectification, Crystal Photoeffect. The Transistor

In the last section we became acquainted with the electronic conductivity of insulating crystals which depends upon the motion of free electrons. These were produced, i.e., raised to the normally unoccupied conduction band, by the absorption of light in the crystal. *In some solids electrons can be raised to the conduction band not only by the absorption of light but also by the thermal vibrations of the lattice, if the temper-*

ature is not too low. This gives rise to a "dark conductivity." These materials are called *semiconductors.*

By definition a semiconductor is a crystal which is an insulator at absolute zero but which, at higher temperatures, has an electronic conductivity that increases exponentially with the temperature. In order to determine whether the conductivity is the result of the migration of electrons and not of ionic diffusion, the *Hall effect* is used. In Fig. 219 an electric field is applied in the y-direction of a semiconductor slab lying in the xy-plane, and a magnetic field acts in the z-direction. Because of their mobility, the electrons moving along the y-direction are then deviated by the magnetic field in the x-direction. The result is a potential difference, the so-called *Hall potential,* perpendicular to the direction of the current. This potential difference can be observed only in the case of electronic conductivity, but not for conduction by ions, because, according to Sec. 7-12, the ion mobility is too small. *The Hall effect thus can be used to distinguish between semiconductors and ionic conductors.*

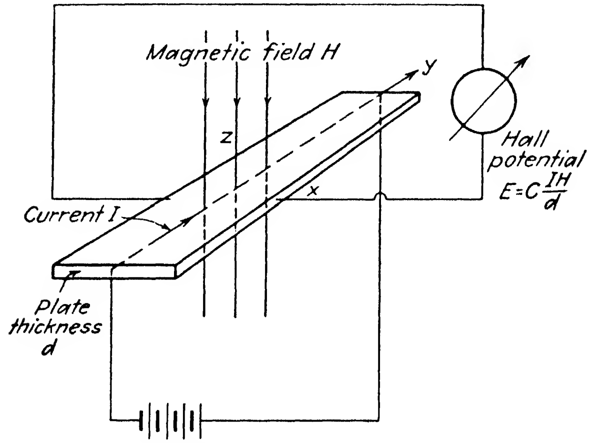


FIG. 219. Schematic sketch of the arrangement for measuring the Hall voltage of semiconductors.

Electronic conductivity in a nonmetallic solid (semiconductor) thus is due to electrons which are released by thermal lattice vibrations (similar to thermal collisions) and raised to the conduction band which is unoccupied at $T = 0$. This is possible in the case of a pure crystal only if the energy difference between the normally empty and the normally filled energy bands is not too large compared with the thermal energy kT, i.e., if it is at most several tenths of a volt in comparison with the 2 to 10 volts in most crystals. Such a crystal with a small gap between the two bands (e.g., germanium with a gap of 0.75 volt) is called an *intrinsic semiconductor.* A crystal which is not an intrinsic semiconductor can be made into a semiconductor (and this is by far the most frequent case) by introducing into it foreign atoms, or a stoichiometric excess of one of the component atoms which has a sufficiently small ionization energy in the crystal. In the energy band model this means that, according to Fig. 220, the energy states D of the valence electrons of these interstitial atoms are situated so closely below the normally unoccupied conduction band B, that electrons can get from D into the band B as a result of the thermal ionization of the metal atoms. This type of crystal is called an *excess semiconductor*, because it requires an excess of metallic ions in the lattice to provide the electrons, or an *N-type semiconductor* because negative charges are released from the D levels. The D levels occasionally are called *donor levels*, because they donate the conduction electrons. The third possible type of semiconductor is called the *P-type or defect semiconductor* because it has a defect of metallic atoms in the lattice or an excess of the electronegative constituents (e.g., oxygen in the metal oxide semiconductors). The energy states of these electronegative atoms lie closely above the band A of the crystal which is completely occupied by electrons (Fig. 221). Only a small amount of energy is necessary to capture an electron, i.e., to bring it from the crystal energy band A to the normally unoccupied levels C of the electronegative atoms, where it is held fast. The "holes" which are thus produced in the previously full band of the crystal have an effective positive charge and are mobile. When an electric field is applied, a current flows which apparently is conducted by positive charges (positive holes, therefore *P-type semiconductor*). The electronegative atoms in the C levels are often called *receptor atoms*, because they receive the electrons from the filled band and thus cause the positive holes.

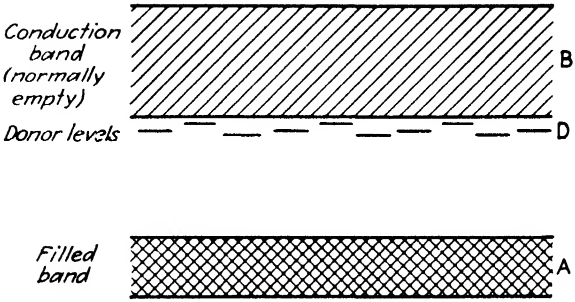


FIG. 220. The energy band diagram for an N-type semiconductor: A = filled energy band; B = empty conduction band; and D = impurity levels of "donor atoms," from which electrons can easily reach the conduction band B.

The excess or defect semiconductors are also called reduction or oxidation semiconductors, respectively, because in the practically important metal oxide semiconductors the excess of the metal or oxygen atoms is produced by chemical reduction or oxidation. Consequently, depending on the preliminary treatment, the same crystal can be an excess or a defect semiconductor, a typical indication for the "structure dependency" of the properties of a semiconductor. Examples of intrinsic semiconductors are pure crystals of silicon and germanium which have recently been extensively investigated. The most important examples of practical excess and defect semiconductors are, respectively, zinc oxide, ZnO, which is heated in vacuum, and cuprous oxide, Cn_2O, heated in oxygen.

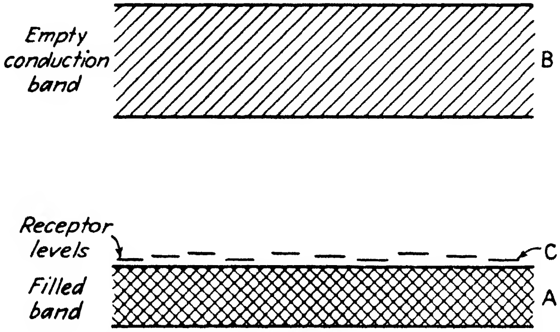


FIG. 221. The energy-band diagram for a *P*-type semiconductor: A = filled energy band; B = empty conduction band; and C = empty impurity levels of "receptor atoms," which can easily be filled from A, thus producing mobile positive holes in A.

Since the energy required for the transition $D \rightarrow B$ (N type) or $A \rightarrow C$ (P type) in the so-called dark conductivity of semiconductors is supplied by thermal collisions, the conductivity can be represented by a formula of the general form

$$\sigma = Ae^{-W/kT} \tag{7-18}$$

in which W is the "activation energy" of the process (the energy difference from D to B or A to C) and is in general, just as is the constant A, a function of the temperature, since W depends on the loosening of the lattice which is determined by the temperature.

On the basis of this theory of semiconductivity, a large number of phenomena of physics, many of which have been known for a very long time, can now be explained. For instance, for many decades, *barium oxide cathodes* have been used in many electronic instruments, because it had been found that at a comparatively low temperature barium oxide has a much higher thermionic electron emission than any metal.

This is also a semiconductor effect. By the "formation" of the barium oxide cathode, an excess of metal atoms is produced in the microcrystals, and these excess metal atoms, being in the D levels of Fig. 220, release electrons into the conduction band B at a sufficiently high temperature. Now the energy necessary to release these electrons from the conductivity band B into free space (effective work function of the semiconductor) is much smaller than that for their release from the normally occupied bands of a metal. This reduced effective work function appears in the exponent in Richardson's equation (7-15) and thus causes, at the same temperature, an increase of the electron emission by many orders of magnitude, as compared with that of a metal.

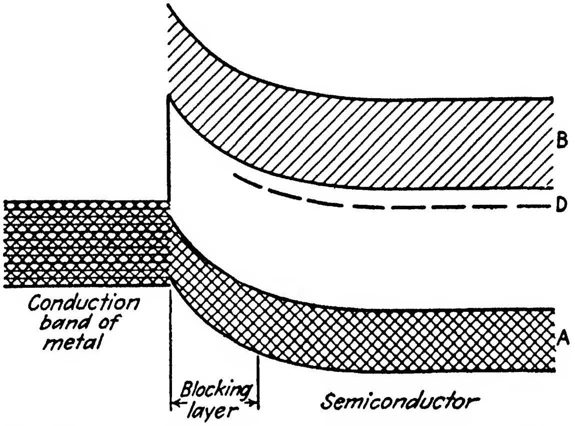


FIG. 222. Energy-band diagram for an N-type semiconductor in contact with a metal; formation of a barrier layer with rectifier properties.

The well known *rectifier and detector properties of numerous metal oxide crystals are also semiconductor phenomena* and can be explained with the aid of the energy band diagram Fig. 220 or 221. If an N-type semiconductor is brought into contact with a metal electrode, electrons from the high impurity levels D "flow down" into the metal, until the electric field set up in the boundary layer causes such a distortion of the energy bands (Fig. 222) that further electron current is impossible, at least without an external field. However, if an alternating electric field is now applied across this metal-semiconductor contact, the energy bands on the right side of Fig. 222 are raised and lowered, alternatingly, with respect to the energy band of the metal electrode at the left. Now with positive polarity of the metal electrode, the bands of the N-type semiconductor are raised to such an extent that electrons from the D levels can again flow down into the metal. When the metal electrode is negative, this is impossible, unless the potential applied to this "blocking layer" becomes too high. In this case the energy bands of the semiconductor at the right side of Fig. 222 are lowered so much that electrons from the metal can penetrate the blocking layer by the tunnel effect (page 234) and flow into the B levels of the semiconductor. Thus this theory requires an easy electron flow from the semiconductor to the metal and a blocking of the current for flow in the opposite direction. However, the resistance of the blocking layer with this polarity should decrease with increasing applied potential. This behavior has been found empirically and holds for most semiconductor rectifiers with the exception that some of them (like cuprous oxide, Cu_2O) are P-type semiconductors, so that we have to use Fig. 221 for our discussion. The detector property of semiconducting crystals is, of course, coupled with rectification and thus need not be treated separately.

According to our above treatment, the decisive phenomenon of N-type semiconductivity is the thermal release of electrons from the impurity levels D. However, the transitions $D \rightarrow B$ or $A \rightarrow C$, which lead to the production of free electrons or free positive holes, respectively, in the semiconductor lattice, need not necessarily be caused by thermal collisions (lattice vibrations), but they can also be caused by light absorption. Consequently, all these semiconductors show weak, longwave absorption bands which can be attributed to these transitions. Absorption of the corresponding radiation by the alkali halides (page 459) produces an electric conductivity which, in contrast to dark conductivity, is called photoconductivity. This release of electrons in the interior of semiconductors (and by shortwave light in other crystals also) is called the "internal photoelectric effect."

Two different methods have been developed to make use of this release of photoelectrons in semiconductors for the detection of radiation. This radiation might cover the whole range from the infrared (3 to 6 μ with PbS, PbSe, and PbTe semiconductors) down to the γ-rays. Either the photoelectric conductivity of the semiconductor is measured directly in a bridge as a function of the incident radiation, or semiconductor photocells are used which themselves produce a photocurrent that is proportional to the number of incident photons. If, in different layers of a semiconducting crystal, very different numbers of electrons can be released photoelectrically, then we get a concentration gradient of free electrons in the interior of the crystal. The crystal acts as an electron-concentration element with an emf which, because of its origin, is called a *photo-emf*. In Cu_2O the photo-emf can amount to as much as 0.4 volt. If a semiconductor is placed between two metal contacts and these are connected to a microammeter, a photocurrent flows when the crystal is irradiated by light: we then have a *semiconductor photoelement.* These have become important for lux meters and other technical applications. The necessary electron concentration gradient can be obtained in two different ways. First, varying concentrations of lattice defects in different layers of the semiconductor can be produced by proper heat treatment, so that varying amounts of electrons are released when the crystal is irradiated. Second, a crystal can be selected which has layers so thick that the incident light will be strongly absorbed. Then many electrons will be released close to the upper surface where the illumination is strong, whereas on the lower side the number of released electrons will be small because only a small portion of the light reaches it. Technically, the first type of photoelement is preferred.

The basic processes of this phenomenon have been carefully studied, in particular for cuprous oxide (a P-type semiconductor) where the necessary C levels (Fig. 221) are due to excess oxygen atoms in the Cu_2O lattice. This is occasionally described as a partial chemical transformation of Cu_2O to CuO by tempering in an oxygen atmosphere. *By this tempering the photoelectric conductivity of pure* Cu_2O *can easily be increased by eight powers of ten!* If a thin layer with only a few excess atoms and thus a high insulation (*barrier layer* or blocking layer) is formed on a crystal layer of good conductivity, and if over these a thin transparent metallic electrode is placed, we have a *semiconductor photoelement or barrier layer photoelement* (Fig. 223). A similar widely used photoelement has been developed using selenium, but little seems to be known about its basic atomic processes.

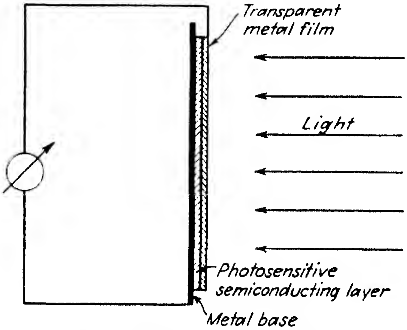


FIG. 223. Diagram of a barrier layer photocell with electric circuit.

We have briefly discussed above the origin of the internal photo-emf. Actually the theoretical problem of the photoelements is much more complicated, because the processes in the barrier layers, which are formed at the contacts between metal and semiconductor, have to be taken into consideration. We have pointed out some of these barrier layer problems above in our discussion of the rectification properties of semiconductors.

We mention finally one of the most recent developments in the field of semiconductivity which combines scientific interest with practical importance. It is the so-called *transistor,* developed in the Bell Telephone Laboratories by Bardeen and Brattain. The transistor is essentially a germanium crystal diode for current amplification. Figure 224 shows the schematic diagram. A small germanium crystal rests on a metal base, whereas two point contacts touch its surface at a distance of 0.05 to 0.25 mm from each other. A current of the order of 1 ma flows in the emitter circuit. As a surface layer of 10^{-5} cm thickness of the crystal is a P-type semiconductor, positive holes are emitted from the positive point contact. *A part of this hole current, which is modulated with the frequency of the signal to be amplified, reaches the negative collector point and changes its current-voltage characteristic.* It is known that in the surface layer of the germanium a barrier layer exists which prevents the flow of electrons from the metal contact into the crystal. *The space charge in the barrier layer under the collector point now is changed by the positive holes from the emitter in such a way that the collector current is increased or reduced with the signal frequency.*

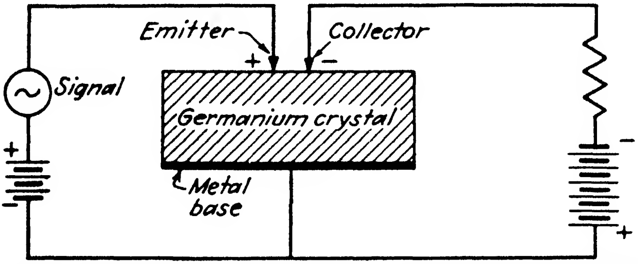


FIG. 224. Electric circuit for the transistor. (*After Bardeen and Brattain.*)

In addition to its probable future role in electronics, the transistor is of interest because its theoretical explanation required a thorough study of the semiconductor properties of germanium which, depending on its treatment, its temperature, and the potential of its metal contacts, shows conductivity due to electrons as well as holes. The electrons and holes either come from the C and D levels, or from transitions between the filled band A or the conduction band B, respectively, and the conduction band of the metal in contact with the semiconductor. For further details the reader is referred to the literature. However, it is clear from this brief discussion that physics of semiconductors has become one of the most important fields of solid state physics.

7-15. Luminescence of Crystal Phosphors

The last important example of processes which depend essentially on the microstructure in the crystal lattice is crystal luminescence, which is closely related to semiconductivity. Many crystals show luminescence after irradiation with shortwave light, α-, β-, or γ-rays, if they have been "*activated*" by addition of a small percentage of metal atoms. The term luminescence here comprises the phenomena of *fluorescence* and *phosphorescence*. Fluorescence is the phenomenon of light emission immediately after radiation has been absorbed in the crystal, i.e., after the normal period of 10^{-8} sec. Phosphorescence, on the other hand, is the more complex process by which light of the same or longer wavelength is emitted for a comparatively long time after irradiation. From these definitions it is apparent that fluorescence and phosphorescence are not always easy to distinguish. Physically, however, *phosphorescence involves a storage of the absorbed energy before its reemission.* Luminescence can be observed in large single crystals: mostly, however, a microcrystalline powder is used.

From what has been said so far, *ideal crystals are not luminescent.* However, certain crystals do show luminescence, particularly at low temperatures, in spite of the fact that they apparently consist of chemically pure homogeneous materials. These crystallize simultaneously in different types of lattices, so that they contain lattice defects and internal surfaces which are essential for phosphorescence. *All other crystals with phosphorescence at normal temperatures consist of a lattice with interstitial or substitutional impurity atoms which serve as emitting centers.* The best known crystal phosphors are zinc and cadmium sulfide, ZnS and CdS, which are used as single or mixed crystals, especially if they are activated by copper or manganese. There is approximately one emitting metal atom for every 10^4 lattice atoms. The oldest phosphors which were investigated carefully by Lenard are the alkaline-earth sulfides and oxides activated by Cn, Mn, Pb, or rare earths. Certain silicates such as Zn_2SiO_4 activated by Mn, fluorspar, CaF_2, activated by heavy metals or rare earths, and numerous tungstates activated by Cr and Mn are also good phosphors. Of especial scientific interest are the alkali halides with heavy metals and rare earths which were investigated by Pohl and Hilsch, because, in contrast to the above-mentioned phosphors, they can be grown in large crystals and therefore can be investigated under more perfect conditions.

The fundamental process of luminescence can be described in a few words. *By absorption of light an electron leaves the highest occupied energy band of the basic lattice and enters a higher unoccupied band (conduction band). It then moves around in the lattice by thermal diffusion until after a shorter or longer time it comes to an activator atom or ion and gives up its energy by emitting light.* In contrast to fluorescence, however, *in phosphorescence we always have a temporary storing of the excitation energy by trapping of the electron in some vacant lattice position or* *similar electron trap.* The dependence of phosphorescence on the structure is obvious from the role of the trapping centers. The temperature dependence is also easy to understand. At low temperatures the trapped electrons have difficulty in getting out of the traps again, because the thermal energy of the lattice vibrations is so small. Only a few electrons are thus released per second, and we have a persistent, but weak, emission which decays exponentially. On the other hand, at higher temperatures the trapped electrons are released at a higher rate from their traps and thus produce a more intensive, but rapidly decaying, phosphorescence. The light sum (the total light emitted during the decay) for the same excitation is the same in both cases, as is to be expected.

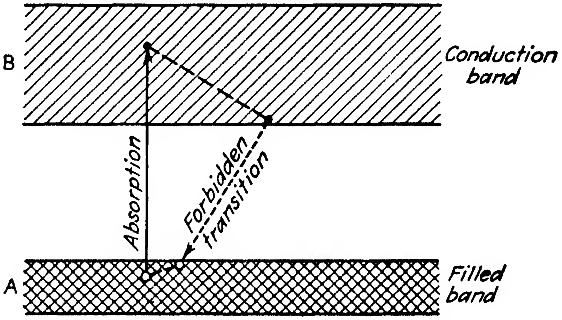


FIG. 225. Energy-band diagram indicating why pure ideal crystals generally do not exhibit any luminescence.

The main features of the great, almost confusing, number of experimental results which experimenters have assembled since the time of Lenard now can be explained. First we ask why an ideal lattice does not exhibit luminescence. This can be explained by means of Fig. 225. If an electron of the occupied energy band A absorbs a photon, it is raised to that state of the higher unoccupied band B which has the same k-quantum number. Now if it does not immediately jump back under emission of light (fluorescence), it will lose energy by interacting with the lattice, until it reaches the lower edge of the energy band B, where it has a different quantum number k. This transfer of energy occurs very rapidly because the period of a lattice vibration with 10^{-14} sec is very short compared with the natural life of 10^{-8} sec in the upper state. The direct reverse transition mentioned above, therefore, is very improbable. Now, however, the electron *cannot* jump back from the lower edge of the upper band into the band A by emitting a photon, because the corresponding quantum state k, to which a transition could occur, is occupied. Luminescence by ideal, pure crystals therefore seems impossible.

The luminescence of metal-activated crystals can be explained by means of the energy band diagram Fig. 226 which is closely related to that of a crystal with N- and P-type semiconductivity. A is again the highest occupied energy band and B the first normally unoccupied conduction band of the normal crystal lattice. In addition, there are again "impurity levels" C and D of substitutional or interstitial foreign atoms. C are the activator states which, in general, are occupied by electrons, because transitions to the completely filled band are impossible. The states D, finally, belong to electron traps, in which electrons can be bound at least temporarily. In the normal case, the D levels are empty. The particular position of the levels C and D with respect to the bands A and B depends on the ionization energy or electron affinity of the atoms under consideration. This explains why certain activator atoms are more favorable than others.

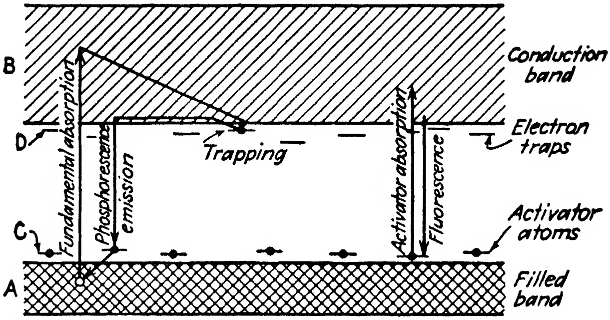


FIG. 226. Energy-band diagram of a crystal with activator atoms and electron traps, for the explanation of crystal phosphorescence.

The mechanism of phosphorescence now seems to be the following: An electron is raised from the filled band A to the unfilled band B by the absorption of sufficiently shortwave light (or impact by α-particles or electrons). In this state it seems to move freely in the lattice. In most crystals this absorption lies in the ultraviolet, since the separation of the energy bands is approximately equal to the first excitation step of the atom. This absorption by the fundamental lattice is very strong because *all* electrons in band A can absorb. The holes resulting from the ionization of electrons in band A are filled by a nonradiative transition from the occupied activator state C. The details of this transition have not yet been explained. The free electron in B now can diffuse to the location of an activator atom and thereby give up energy to the lattice, until it makes a transition to the state C under emission of light. On the other hand, it may, before it meets an activator atom, be trapped in a state D. We speak of trapping, for the transition $D \rightarrow A$ is not possible because the gaps in A are filled from C, and the transition $D \rightarrow C$ is not possible because the separation of the two places in general is too great. Thus *the electron and its excitation energy are stored in D until, by transfer of thermal energy or by the absorption of longwave radiation, it gets back to state B, diffuses to an activator atom, and emits light by a transition to C. This seems to be the essential mechanism of phosphorescence which thus clearly depends on the number of electron traps (states D), the number of emitting atoms (states C), and the temperature.* The very longwave absorption $D \rightarrow B$ is very weak because of the relatively few occupied D states; by measuring the corresponding absorption, the number of occupied D states can be determined. The trapping states D, which are necessary for phosphorescence, can also be reached by direct absorption from the fundamental band A. This absorption is very much weaker than the fundamental lattice absorption $A \rightarrow B$, but stronger than the longwave absorption $D \rightarrow B$, which depends on the number of *occupied* D states. There is, finally, the possibility of absorption by the activator atoms themselves (transition $C \rightarrow B$). This absorption is proportional to the number of activator atoms (C states).

It is clear that the diffusion of the liberated electrons in band B represents an electric convection current and therefore, if an electric field is applied to the crystal, it can be measured electrically just as in the corresponding cases of the F electrons (page 459) or the photoelectrons in semiconductors.

No doubt, the theoretical picture described here is much too simple. It will require considerable development and refinement before it can explain in detail the mass of existing empirical results which is continuously increasing at a surprising rate.

Recently, for instance, phosphors with two kinds of activator atoms (e.g., SrSe with Cu and Sm) were investigated in more detail. While the copper atom as the "*dominant activator*" is responsible for the emission process, the *auxiliary activator* (in this case samarium) supplied the trapping states (D in Fig. 226). Leverenz, furthermore, distinguishes between s and i phosphors, according to whether the activator occupies an interstitial place or, as a substitutional atom, a normal lattice place. Many known differences in the behavior of different phosphors, for instance with respect to their photoconductivity, can be understood on the basis of this distinction.

We mention in concluding, that the macrocrystalline structure of phosphors seems to be responsible for their different behavior with respect to excitation by light or particles. High light efficiency from corpuscular excitation is obtained only with larger, reasonably perfect crystals in which the free electrons travel a sufficiently long mean free path in the lattice until they find an activator atom. *Light* absorption with subsequent emission, on the other hand, occurs very frequently either by the activator atoms themselves (transition $C \rightarrow B$) or in the close neighborhood of an activator atom, so that a good phosphorescence yield with light excitation is found in crystal powders in which the electrons cannot migrate very far.

7-16. Collision Processes at the Surface of Solids

To conclude our atomistic treatment of solid state physics, we shall briefly consider a number of interesting collision processes which occur at the surface of solids. First, there is the question of the possible ionization of atoms at the surface of an incandescent metal. Second, there are the phenomena connected with an energy transfer to the crystal lattice or crystal electrons in collisions between electrons, ions, normal or excited atoms on the one side, and the surface of a solid body on the other.

In discussing the ionization at an incandescent surface we start from the experimental results according to which alkali atoms in the vapor phase that impinge on an incandescent tungsten strip leave the strip as ions, if they are atoms of the heavy alkalies potassium, rubidium, and cesium. However, atoms of the lighter alkalies lithium and sodium leave the incandescent tungsten surface unionized. On the other hand, *all* alkali atoms are ionized at an incandescent platinum sheet. We may assume that the atoms which come in contact with the metal surface are at first adsorbed and then, as a result of the high temperature of the sheet, evaporate. Then it is only a question of whether a departing atom carries its outer electron with it or leaves it in the metal.

We can understand the whole situation if we imagine that the evaporation may occur in several steps (analogous to the Born cycle, page 429). Instead of the neutral atom we can imagine that first an ion evaporates. By adding energy equal to the work function Φ we then release an electron from the metal and let it recombine with the ion to form an atom, thereby releasing the ionization energy E_i. Actually the process occurs in such a manner that the least possible energy is required.

Thus, if

$$\Phi > E_i \tag{7-19}$$

the electron remains in the metal and the atom leaves the surface as an ion. For in order to liberate an electron from the metal and let it recombine with the ion, the energy $\Phi - E_i$ would have to be expended. In the normal case, however, where

$$E_i > \Phi \tag{7-20}$$

the energy $E_i - \Phi$ is gained if an electron is "pulled" out of the metal and united with the ion. In this case the atoms will *not* be ionized in collisions with the incandescent metal. These theoretical conclusions are in good agreement with the experimental results. The heavy alkali atoms K, Rb, and Cs have ionization potentials which are smaller than the work function of tungsten, 4.5 ev. The light alkalies Li and Na, on the other hand, have larger ionization potentials. However, the ionization potentials of *all* alkalies lie below the work function of platinum with 6.2 ev; thus *all* alkali atoms are ionized by incandescent platinum.

The reverse process occurs at the cathode of glow and arc discharges. Following the lines of force of the electric field, positive gas ions from the discharge arrive at the cathode and are neutralized. This implies that they pull an electron out of the cathode and leave it as an atom. According to our previous discussion, this neutralization is energetically possible only if (7-20) is fulfilled. This, however, is the case, if we disregard discharges in the heaviest alkali vapors and a few electrode metals with particularly high work functions.

Now we shall discuss the mechanism of the energy transfer to the metal. For in our example the energy E_i is released in the neutralization of the arriving ion and is used partly for the release of a metal electron (under expenditure of the work function Φ), whereas the rest, $E_i - \Phi$, heats the cathode. This mechanism of neutralization may be regarded in a very pictorial manner. The positive ion attracts electrostatically a quasi-free metal electron and recombines with it. The process is more difficult to understand, if the condition

$$E_i > 2\Phi \tag{7-21}$$

is fulfilled. In this case the potential energy released in the neutralization of the ion is larger than the work function of *two* electrons. In these cases, actually two electrons are released. This secondary electron emission by ion bombardment is observed particularly in discharges in the noble gases helium and neon with their high ionization potentials. However, a metastable helium atom (see page 139) can also give up its potential energy (excitation energy) in a collision of the second kind with a metal surface and release one or two secondary electrons. The mechanism of this process is difficult to understand because no electrostatic effect is possible now, and the purely mechanical effect of the collision does not explain the emergence of an electron (or even two), because the impact is directed toward the interior of the metal. Basically the same difficulty is met in the transfer of purely kinetic energy to the metal in collisions. Fast electrons of sufficient kinetic energy, in collisions with metals, can release a large number of secondary electrons, a technically very important process. Fast positive ions, similarly, can liberate, though with much smaller efficiency, electrons at the cathode of glow discharges, and furthermore cause the phenomenon of cathode sputtering, i.e., remove whole atoms from the metal surface.

It seems very probable that thermal processes and not simply mechanical collision effects are responsible for all these phenomena. We know, for example, that the anode of a high-current carbon arc and of certain spark discharges can be heated to violent evaporation by the discharge electrons which release on the anode their kinetic energy as well as the energy corresponding to the work function. The possibility has been considered that cathode sputtering is due to a similar effect of positive ions. However, this explanation seemed to be ruled out by the observation that the heat conductivity seems to have no influence on cathode sputtering. To cite one example: metallic silver and the poorly conducting silver chloride both exhibit equally strong cathode sputtering. However, all these difficulties are removed if, according to Seeliger, these phenomena are regarded as a very local overheating or, to use a familiar expression, as an extreme case of local lattice loosening. We have to imagine, then, that the temperature of a very small region of the cathode exceeds by far the critical temperature so that we have a gas-like state of matter of high density. Whether there exists an actual thermal equilibrium between the original lattice particles and the colliding particles seems more than doubtful, however.

In concluding this chapter on atomistic solid state physics, we again stress the fact that a satisfactory and rigorous presentation of this field, according to the present state of our knowledge, is not yet possible. However, we hope that our discussion of the most important problems of solid state physics has, on the one side, shown how fruitful the atomistic method of treatment has proved to be and that it may, on the other side, have helped us to realize how important it is to pay close attention in the future to this promising field of atomic physics.

Literature

General Reference Books

Vol. 24 of the Handbuch der Physik, Struktur der zusammenhängenden Materie, 2d ed., Springer-Verlag, Berlin, 1934.

Seitz, F.: The Modern Theory of Solids, McGraw-Hill Book Company, Inc., New York, 1940.

SECTION 1:

Frenkel, J.: Kinetic Theory of Liquids, Oxford University Press, New York, 1946.

Gingrich, N. S.: The diffraction of X-rays by liquid elements, *Rev. Mod. Phys.*, **15**, 90 (1943).

Lehmann, O.: Die Lehre von den flüssigen Kristallen, J. F. Lehmann, Wiesbaden, 1918.

SECTION 3:

Finkelnburg, W.: Kontinuierliche Spektren, Springer-Verlag, Berlin, 1938.

SECTION 4:

Bragg, W. H., and W. L. Bragg: The Crystalline State, George Bell & Sons, London, 1933.
Internationale Tabellen zur Bestimmung von Kristallstrukturen, 2d ed., Verlagsbuchhandlung Gebrüder Bornträger, Berlin, 1944.

SECTION 5:

Fermi, E.: Moleküle und Kristalle, J. A. Barth, Leipzig, 1938.

SECTION 6:

Barrett, C. S.: Structure of Metals, McGraw-Hill Book Company, Inc., New York, 1943.
Boas, W.: An Introduction to the Physics of Metals and Alloys, John Wiley & Sons, Inc., New York, 1948.
Dehlinger, U.: Chemische Physik der Metalle und Legierungen, Akademische Verlagsgesellschaft m.b.H., Leipzig, 1939.
Hume-Rothery, W.: Structure of Metals and Alloys, New York Chemical Co., New York, 1939.
Mott, N. F., and H. Jones: The Structure of Metals and Alloys, Oxford University Press, New York, 1936.
Schmid, E., and W. Boas: Kristallplastizität, Springer-Verlag, Berlin, 1935.
Seitz, F.: The Physics of Metals, McGraw-Hill Book Company, Inc., New York, 1943.
Wilson, A. H.: The Theory of Metals, Cambridge University Press, London, 1936.
Zener, C.: Elasticity and Anelasticity of Metals, Chicago University Press, Chicago, 1948.

SECTION 7:

Brillouin, L.: Wave Propagation in Periodic Structures, McGraw-Hill Book Company, Inc., New York, 1946.

SECTIONS 8 AND 9:

Fröhlich, H.: Elektronentheorie der Metalle, Springer-Verlag, Berlin, 1936.
Justi, E.: Leitfähigkeit und Leitungsmechanismus fester Stoffe, Vandenhoeck und Ruprecht, Göttingen, 1948.
Mott, N. F., and R. W. Gurney: Electronic Processes in Ionic Crystals, Oxford University Press, New York, 1940.
Wilson, A. H.: Semiconductors and Metals, Cambridge University Press, London, 1939.

SECTION 11:

Becker, R., and W. Döring: Ferromagnetismus, Springer-Verlag, Berlin, 1939.
Stoner, E. C.: Ferromagnetism, *Rep. Progr. Phys.* **11,** 43 (1948).

SECTION 12:

Jost, W.: Diffusion und chemische Reaktion in festen Stoffen, T. Steinkopf, Leipzig, 1937.

SECTION 13:

von Angerer, E.: Wissenschaftliche Photographie, Akademische Verlagsgesellschaft, Leipzig, 1931.

Mott, N. F., and R. W. Gurney: Electronic Processes in Ionic Crystals, Oxford University Press, New York, 1940.

Pohl, R. W.: "Optik" aus "Einführung in die Physik", Springer-Verlag, Berlin, 1948.

Seitz, F.: Color centers in alkali-halide crystals, *Rev. Mod. Phys.* **18**, 384 (1946).

SECTION 14:

Gudden, B.: Lichtelektrische Erscheinungen, Springer-Verlag, Berlin, 1928.

Lange, B.: Die Photoelemente und ihre Anwendung, J. A. Barth, Leipzig, 1940.

Mott, N. F., and R. W. Gurney: Electronic Processes in Ionic Crystals, Oxford University Press, New York, 1940.

Wilson, A. H.: Semiconductors and Metals, Cambridge University Press, London, 1939.

SECTION 15:

Kroeger, F. A.: Some Aspects of the Luminescence of Solids, Elsevier Publ. Co., New York and Amsterdam, 1948.

Pringsheim, P.: Fluorescence and Phosphorescence, Interscience Publishers, New York, 1949.

Pringsheim, P., and M. Vogel: Luminescence of Liquids and Solids, Interscience Publishers, New York, 1946.

Riehl, N.: Physikalische und Technische Anwendungen der Lumineszenz, Springer-Verlag, Berlin, 1941.

Seitz, F., *et al.*: Solid Luminescent Materials, John Wiley & Sons, Inc., New York, 1948.

APPENDIX

The Most Important Constants and Energy Relations of Atomic Physics

Planck's constant	h	$= (6.623 \pm 0.001) \times 10^{-27}$ erg-sec
Velocity of light	c	$= (2.99793 \pm 0.00001) \times 10^{10}$ cm sec^{-1}
Charge of the electron	e	$= (4.8024 \pm 0.0005) \times 10^{-10}$ esu $= (1.60199 \pm 0.00016) \times 10^{-19}$ coulombs (amp-sec)
Rest mass of the electron	m_0	$= (9.1055 \pm 0.0012) \times 10^{-28}$ gm
Rest mass of the proton	M_p	$= (1.6723 \pm 0.0001) \times 10^{-24}$ gm
Rest mass of the neutron	M_N	$= (1.6747 \pm 0.0002) \times 10^{-24}$ gm
Atomic weight of the proton	A_p	$= 1.007597 \pm 0.000003$
Atomic weight of the neutron in physical units	A_N	$= 1.008987 \pm 0.000003$
Atomic weight of the electron	A_e	$= (5.4862 \pm 0.0006) \times 10^{-4}$
Avogadro's number	N_0	$= (6.0235 \pm 0.0004) \times 10^{23}$
Faraday	F	$= 96522 \pm 7$ coulombs/mole
Boltzmann's constant	k	$= (1.3803 \pm 0.0001) \times 10^{-16}$ erg deg^{-1}

Relations between different energy units used in atomic physics:

$$1 \text{ ev} = 8{,}067.5 \text{ cm}^{-1} = 1.602 \times 10^{-12} \text{ erg} = 23.04 \text{ kcal/mole}$$

$$1 \text{ atomic mass unit} = 931 \text{ Me}$$

Relation between excitation potential E (electron volts) and wavelength λ in angstroms of corresponding radiation:

$$\lambda(\text{A}) \times \text{E (ev)} = 12{,}400$$

INDEX