UNIVERSAL
LIBRARY

OU_158459

UNIVERSAL
LIBRARY

THE ELEMENTS OF QUANTUM MECHANICS

THE ELEMENTS OF QUANTUM MECHANICS

BY

SAUL DUSHMAN, Ph.D.

Assistant Director, Research Laboratory, General Electric Co., Schenectady, N. Y.

NEW YORK

JOHN WILEY & SONS, INC.

LONDON: CHAPMAN & HALL, LIMITED

1938

COPYRIGHT, 1938
BY SAUL DUSHMAN

All Rights Reserved
This book or any part thereof must not be reproduced in any form without the written permission of the publisher.

Printed in U. S. A.

Printing
F. H. GILSON CO.
BOSTON

Composition
TECHNICAL COMPOSITION CO.
BOSTON

Binding
STANHOPE BINDERY
BOSTON

DEDICATED TO
A. L. D.

PREFACE

In the summer of 1932 the author was invited by Professor W. Lloyd Evans, Chairman of the Department of Chemistry, Ohio State University, Columbus, Ohio, to give a series of lectures on quantum mechanics. For the opportunity thus afforded him for study of this subject in a university atmosphere the author wishes to express his gratitude to Professor Evans.

The notes prepared for these lectures were subsequently published as a serial in the *Journal of Chemical Education* (May 1935 to August 1936, inclusive). To the editor, Dr. Otto Reinmuth, the author is indebted for many helpful suggestions with regard to method of presentation and contents.

Since no reprints of the series were made available it was suggested that the contents be revised for publication in book form. As stated in the first chapter, the writer's aim has been to present the subject in such a manner that its essential concepts and logic may be readily comprehended by those who have not had any intensive training in mathematics beyond calculus. For this reason there has been presented in a number of cases a great deal more detail of the mathematical development than would seem necessary to those readers who are familiar with more advanced branches of mathematics.

The author lays no claim to being an expert in the field of quantum mechanics. But, like many other workers in science, he has felt a strong desire to learn something about its technic and applications. This volume may therefore be regarded in a sense as a series of notes which have served to clarify, at least to his own satisfaction, some of the difficulties which he, together with probably a considerable number of other students, has encountered in attempting to understand the subject. Should the contents of this volume prove of any assistance to others in enabling them to proceed with the study of more advanced treatises, he will feel amply rewarded for a task which has indeed been a source of intellectual pleasure.

He also wishes to take this opportunity of expressing his appreciation of the sympathetic support of Dr. W. D. Coolidge, the Director of the Research Laboratory of the General Electric Company, in a task which could scarcely have been carried through without it.

To Dr. Frederick Seitz of this Laboratory the author is deeply grateful both for helpful discussion and clarification of difficult topics and for reading a considerable part of the galley proof.

Finally the author wishes to express his indebtedness to Miss Elizabeth Gage for invaluable assistance in the typing of the manuscript and in the even more tedious task of proof-reading.

SAUL DUSHMAN.

RESEARCH LABORATORY
GENERAL ELECTRIC COMPANY
Schenectady, New York
October 7, 1937.

CONTENTS

CHAPTER

I QUANTUM PHENOMENA

1 Introductory Remarks, 1. 2 Energy States of Atomic Systems, 4. 3 Resonance Radiation, 7. 4 Photoelectric Effect, 8. 5 The Compton Effect, 10. 6 Undulatory Phenomena Associated with Corpuscles, 11. 7 Principle of Indeterminacy, 12.

II THE SCHROEDINGER EQUATION IN ONE DIMENSION

1 The Concepts of Quantum Mechanics, 19. 2 Some Fundamental Differential Equations, 20. 3 Equations for the Propagation of Wave Motion, 29. 4 Differential Equation for the Vibration of a String, 31. 5 S. Equation for One Coördinate, 36. 6 Motion of Electrons in Absence of Field of Force, 39. 7 Operator Method of Deriving the S. Equation, 42.

III PROBLEMS OF POTENTIAL BARRIERS

1 Reflection of Electrons at Semi-Infinite Potential Barriers, 47. 2 Electrons between Semi-Infinite Potential Barriers, 55. 3 Transmission of Electrons through Potential Barrier of Finite Extent, 63. 4 Application to Radioactive Disintegration, 70. 5 Emission of Electrons in Presence of Intense Fields, 72.

IV CLASSICAL THEORY OF ATOMIC DYNAMICS

1 Relation between Quantum Mechanics and Classical Dynamics, 77. 2 Kinetic and Potential Energy, 77. 3 Generalized Coördinates, 82. 4 Lagrange's Equations, 86. 5 Motion of Electron in Bohr Atom in Absence of Quantizing Conditions, 89. 6 Canonical Equations, 94. 7 Hamilton's Principle and the Principle of Least Action, 96. 8 The Wilson-Sommerfeld Quantum Conditions, 98. 9 Mean Values for Central Orbits, 103. Supplementary Note 1. Equation of Ellipse in Polar Coördinates, 106. Note 2. Relation between Average Kinetic and Potential Energies for Central Orbits (Virial Theorem), 109.

V THE LINEAR HARMONIC OSCILLATOR

1 Formulation of the S. Equation and Solution, 111. 2 Distribution Functions for Harmonic Oscillator, 122. 3 Determination of Average Values, 125. 4 Probabilities of Transition, 131. Supplementary Note 1. Proof That Solutions of the S. Equation for One Degree of Freedom Are Orthogonal, 133. Note 2. The Gauss Error Function, 136.

CHAPTER

VI THE RIGID ROTATOR

1 The S. Equation in Three Dimensions, 137. 2 Rules for Transformation of Coördinates, 138. 3 The Rigid Rotator with Fixed Axis, 143. 4 The Rigid Rotator with Free Axis, 146. 5 Legendre Equation of Order Zero, 149. 6 Associated Legendre Functions, 154. 7 Geometrical Interpretation of Surface Spherical Harmonics, 158. 8 The Physical Significance of the Characteristic Functions, 160. 9 Angular Momentum for Motion of Rotator, 166. Supplementary Note 1. Expansion of an Arbitrary Function in Terms of an Orthogonal System of Functions, 171.

VII THE HYDROGEN ATOM

1 Bohr Theory of the Hydrogen-Like Atom, 178. 2 Hydrogen Atom as a Potential Barrier Problem, 179. 3 The S. Equation for the Hydrogen-Like Atom, 180. 4 The Radial S. Equation. Laguerre Functions, 182. 5 The Complete Solution of the S. Equation for the Hydrogen Atom, 189. 6 Quantum Mechanics Interpretation of the Characteristic Functions, 192. 7 Comparison between Deductions from Classical and Quantum Mechanics, 197. 8 Relation between Kinetic and Potential Energy Deduced from S. Equation, 201. 9 Angular Momentum of Electron, 206.

VIII VAN DER WAALS FORCES

1 Van der Waals' Equation, 208. 2 Two Interacting Pendulums, 211. 3 Electron in Two Adjacent "Boxes," 214. 4 Coupled Linear Harmonic Oscillators, 219. 5 Van der Waals Interaction Energy for Two Hydrogen Atoms, 228. 6 Van der Waals Energy for Other Atoms and Derivation of the Constant "a," 233. 7 Energies of Evaporation and Sublimation, 236.

IX PERTURBATION THEORY

1 Introductory Remarks, 239. 2 Perturbation Theory — First-Order Terms, 241. 3 Second-Order Perturbation Terms, 245. 4 Perturbation Terms in Relation to Matrix Elements, 247. 5 Perturbation Theory for Degenerate Case, 249. 6 Application to Coupled Linear Oscillators, 252. 7 Interaction Energy of Two Hydrogen Atoms, 256. 8 Application of Perturbation Theory to Stark and Zeeman Effects, 259.

X THE HELIUM ATOM — PERTURBATION METHOD

1 Atomic Units, 262. 2 The S. Equation for the Helium Atom, 264. 3 The Normal State of the Helium Atom, 265. 4 Excited States of the Helium Atom, 268. 5 Calculation of Matrix Elements V_{11} and V_{12}, 272. 6 Exchange Energy, 274. 7 Orthohelium and Parhelium, 276. 8 Electron Spin. Pauli's Exclusion Principle, 278. 9 Multiplet Levels in Spectrum of Helium, 280.

XI THE HELIUM ATOM — VARIATIONAL METHOD

1 Minimum Energy Principle. The Variation Theorem, 283. 2 The Energy for the Normal State of the Helium Atom, 289. 3 Variational Method of Hylleraas, 293. 4 The Problem of the Many-Electron Atom, 300. 5 The Method of the Self-Consistent Field, 303. 6 Slater's Treatment of the Many-Electron Problem, 304.

CHAPTER

XII THE HYDROGEN MOLECULE

1 Heitler-London Method, 310. 2 Physical Interpretation, 315. 3 Application of the Pauli Exclusion Principle, 319. 4 Calculation of Perturbation Energy Terms, 320. 5 The Variational Method, 326. 6 Method of Molecular Orbitals, 329. 7 The Dissociation Energy of H_2, 332. Supplementary Note 1. Transformation from Cartesian to Confocal Elliptic Coördinates, 336.

XIII VIBRATIONAL AND ROTATIONAL STATES OF THE HYDROGEN MOLECULE

1 General Remarks, 340. 2 Vibrational Energy States, 343. 3 Rotational Energy Levels, 349. 4 Orthohydrogen and Parahydrogen, 352.

XIV VALENCE BONDS; ACTIVATION AND RESONANCE ENERGY

1 The Homopolar Bond, 356. 2 Interaction of Two H_2 Molecules, 357. 3 Pauling's Theory of Directed Valence Bonds, 358. 4 Change in Quantization of Bond Eigenfunctions, 361. 5 Slater's Treatment of Polyatomic Molecules, 364. 6 Method of Localized Pairs, 374. 7 Activation Energy, 376. 8 Resonance Energy, 381.

XV QUANTUM MECHANICS THEORY OF RADIATION

1 Classical Theory of Radiation, 385. 2 S. Equation Involving Time, 388. 3 Quantum Mechanics Expressions for Dipole Moment and Intensity of Radiation, 392. 4 Selection Rule for Linear Harmonic Oscillator, 393. 5 Selection Rules for Quantum Numbers m and l, 394. 6 Wave Packets. Group Velocity, 398. 7 Relation between Eigenfunction Representation and Particle Velocity, 402. 8 Spreading Out of Wave Packet in Time, 405.

APPENDIX

I TREATISES AND REVIEW PUBLICATIONS ON QUANTUM MECHANICS 409

II VALUES OF PHYSICAL CONSTANTS 411

III SPECIAL TABLES OF MATHEMATICAL FORMULAS 413

IV SOME FUNDAMENTAL THEOREMS AND DIFFERENTIAL EQUATIONS 419

INDEX 439

ELEMENTS OF QUANTUM MECHANICS

SAUL DUSHMAN

ERRATA

NOTE: The author will appreciate having his attention drawn to any errata in the text other than those mentioned in the following list.

Page	
38	In the second equation from the bottom insert minus sign before the right-hand side.
49	In equation (8) replace α in exponent by β.
66	The equation $A\overline{A} + B\overline{B} = 1$ should read $A\overline{A} - B\overline{B} = 1$.
67	In the last two equations at the bottom of the page replace $\frac{1}{\beta}$ by $\frac{2}{\beta}$.
68	In the table at the top of the page the heading of the first column should read $2a \times 10^8$.
84 and 85	In the footnotes Appendix III should read Appendix IV.
87	In equation (31) replace the factor x_i by $\dot{x}_i$.
115	In the ninth line from the bottom the lower limit of integration should read 0.
116	Near the top of the page the second line of the first equation should read $-2 \cdot 2a_{n-2}$.
125	In the equation near the top of the page insert the factor c^2 in the right-hand side.
142	In equation (7b) change the coefficient $\frac{1}{r}$ in the last term to $\frac{2}{r}$.
154	The last expression in the last equation should read $(1 - x^2)^{\frac{m}{2}+1} \cdot \frac{d^2Y}{dx^2}$.
187	In the exponent of ϵ replace $\frac{x^2}{2}$ by $\frac{x}{2}$ in the three equations in which it occurs.
225	Equation (23) should read $\frac{d^2\phi}{dz^2} + a\left(E + \frac{e^2F^2}{2k} - \frac{kz^2}{2}\right)\phi = 0.$
228	Replace z_1 and z_2 by q_1 and q_2, respectively, in the equation for V, and similarly in lines 7, 8 and 9.
231	In the right-hand side of equation (37b) insert the factor a_0^6.
234	Replace dyne · cm.^{6} by dyne · cm.^{7}.
235	In equation (46), insert the factor $\frac{4}{3}$ in the middle term.
389	In equation (16) replace $\nabla\psi$ by $V\psi$.

ILLUSTRATIONS

		PAGE
FIG. 1	Energy levels and lines in arc spectrum of sodium	6
FIG. 2	Illustrating the theory of the Compton effect	10
FIG. 3	Illustrating the Principle of Indeterminacy	14
FIG. 4	Geometrical interpretation of a simple differential equation and of its solution	21
FIG. 5	Illustrating the superposition principle of wave motions	25
FIG. 6	Plot of hyperbolic cosine and sine functions	28
FIG. 7	Illustrating derivation of partial differential equation for vibration of a string	32
FIG. 8	Semi-infinite potential barrier; Case $E > U_0$	47
FIG. 9	Semi-infinite potential barrier; Case $E < U_0$	47
FIG. 10	Coefficients of reflection R and transmission T for electrons as a function of E/U_0	50
FIG. 11	Form of eigenfunction for penetration in region $E < U$	54
FIG. 12	Illustrating behavior of electron in potential "box"	55
FIG. 13	Symmetrical eigenfunctions for electron in "box"	57
FIG. 14	Antisymmetrical eigenfunctions for electron in "box"	57
FIG. 15	Illustrating penetration of particle through rectangular potential barrier	63
FIG. 16	Illustrating penetration of particle through barrier of arbitrary shape	68
FIG. 17	Potential energy curve for alpha particle in nucleus of radioactive atom	71
FIG. 18	Illustrating "cold-cathode" emission of electrons	73
FIG. 19	Potential distribution at surface of metal covered with monatomic film of more electropositive element	75
FIG. 20	Illustrating equation (16)	82
FIG. 21	Illustrating cylindrical coördinates	82
FIG. 22*a*	Illustrating relation between rectangular and spherical coördinates	83
FIG. 22*b*	Projection on meridian plane	84
FIG. 22*c*	Projection on equatorial plane	84
FIG. 23	Illustrating derivation of equation for ellipse in polar coördinates	106
FIG. 24	Eigenfunctions for lower energy states of linear harmonic oscillator	120
FIG. 25	Probability of occurrence of oscillating particle according to quantum mechanics (Curve A) and according to classical theory (Curve B)	124
FIG. 26	Plots of the first four Legendre polynomials as functions of θ	151
FIG. 27	Plots of Legendre polynomials P_4–P_7 inclusive	151
FIG. 28	Plots of the normalized Legendre functions P_3^m (cos θ)	158
FIG. 29	Geometrical illustration of zonal harmonics	159
FIG. 30	Geometrical illustration of tesseral harmonics	159
FIG. 31	Geometrical illustration of sectorial harmonics	160
FIG. 32	Plots of the zonal distribution functions corresponding to $k = 3$	163
FIG. 33	Plots of the zonal distribution functions corresponding to $k = 3$	164

PAGE

FIG. 34 Probability density function corresponding to associated Legendre function $P_3^0(\cos\theta)$ 165
FIG. 35 Probability density function corresponding to associated Legendre function $P_3^3(\cos\theta)$ 165
FIG. 36 Illustrating Fourier's series analysis 172
FIG. 37 Plots of normalized radial functions for different electronic states of hydrogen atom 189
FIG. 38 Angular distribution functions for different electronic states of hydrogen atom 193
FIG. 39 Radial distribution function (D) for different electronic states of hydrogen atom, and Bohr orbits for comparison 195
FIG. 40 Plots of characteristic function (ϕ_{100}), density per unit length (ρ), and distribution function (D) for normal state of hydrogen atom 196
FIG. 41 Illustration of probability of occurrence of electron for excited state of hydrogen atom ($n = 2, l = 0$) 197
FIG. 42 The charge density distribution for $n = 2, l = 1, m = 0$ 197
FIG. 43 Illustrating the two modes of vibration of two interacting pendulums .. 211
FIG. 44 Motion of each pendulum as a superposition of the two modes 211
FIG. 45 Plot of the function $y = \sin 2\pi\nu t \cdot \sin 2\pi(\nu t/9)$ 213
FIG. 46 Illustrating the two eigenfunctions for electrons in two adjacent "boxes" 215
FIG. 47 Illustrating the two eigenfunctions for electron in ionized hydrogen molecule (H_2^+) 218
FIG. 48 Interaction of two dipoles arranged in the same direction 220
FIG. 49 Phase relations for two coupled linear harmonic oscillators 228
FIG. 50 Illustrating van der Waals interaction of two hydrogen atoms 229
FIG. 51 Van der Waals potential energy curves for several gases 235
FIG. 52 Symmetry and antisymmetry of electrons 275
FIG. 53 Energy levels in spectrum of helium 277
FIG. 54 Illustrating notation used in formulating potential energy function for interaction of two hydrogen atoms 310
FIG. 55 Energy curves for the two modes of interaction of two hydrogen atoms; curve S corresponds to the symmetric, and curve A to the antisymmetric mode 315
FIG. 56 Plots of the total energy, Coulomb energy and exchange energy as functions of internuclear distance, for the two modes of interaction of two hydrogen atoms 316
FIG. 57 Relation between the terms in the expressions for the energy of interaction of two hydrogen atoms 317
FIG. 58 Electron distribution for elastic reflection of two hydrogen atoms ... 318
FIG. 59 Electron distribution for hydrogen molecule formation 318
FIG. 60 Plot of the function S^2 as a function of internuclear distance 322
FIG. 61 Confocal elliptic coördinates 336
FIG. 62 Potential energy function and vibrational energy levels for normal state of H_2 molecule 341
FIG. 63 Normal and first excited electronic energy levels for H_2 molecule and associated vibrational energy levels 342
FIG. 64 Illustrating transitions between rotational energy levels associated with two different vibrational energy levels for H_2 molecule 343
FIG. 65 Illustrating the vibrational motion of two nuclei 344

PAGE

FIG. 66 Plot of ΔG_v and internuclear distance (r_0) as functions of the wave numbers of the vibrational levels for the normal state of H_2 347

FIG. 67 Plot of B_v versus v for the normal ($1s^1\Sigma$) and two excited electronic states of H_2 351

FIG. 68 s and p eigenfunctions (Pauling) 360

FIG. 69 Tetrahedral sp eigenfunction (Pauling) 363

FIG. 70 Interaction energies of four atoms 375

FIG. 71 Energy surface (showing contour lines) for three atoms arranged linearly 378

FIG. 72 Relation between energies of activation and of reaction for reversible reaction 379

FIG. 73 Potential energy (Morse) curves for H_2, HBr and Br_2 380

FIG. 74 (*a*) Suggested valence bond structures for benzene ring 383

(*b*) Five canonical valence-bond structures for benzene, leading to resonance phenomenon 383

THE ELEMENTS OF QUANTUM MECHANICS

CHAPTER I

QUANTUM PHENOMENA

1.1 Introductory Remarks. A little over a third of a century has passed since Lord Kelvin, in an address before the British Association, pointed out that there were apparently two clouds upon the scientific horizon. One of these was represented by the experiment of Michelson and Morley; the other involved the failure of classical theory in accounting for the observations on the energy distribution in the radiation emitted by a black body. The first difficulty led Einstein to formulate his special theory of relativity, and, subsequently, a more generalized form of the theory which involved a radical interpretation of the force of gravity. The second difficulty led Planck to formulate a theory of energy quanta which, through the work of Einstein, A. H. Compton, N. Bohr, and others, has led to a corpuscular theory of the interaction of matter and radiation. This quantum theory entered upon a second phase in 1926 with the discovery of the undulatory nature of corpuscular motion, and, through the theoretical investigations of Heisenberg, Schroedinger, and Dirac, there has been developed a totally new point of view on the nature and behavior of electrons, atoms, and molecules.

The system of concepts and mathematical technic originated by these investigators, together with the applications of these new methods to physical and chemical problems, constitute what has been designated as the new Quantum Mechanics.

Although a more comprehensive discussion of this theory requires a considerable knowledge of mathematical technic with which most chemists and many physicists are not familiar, the writer believes that the essential features of the new point of view may be presented without recourse to such highly intricate mathematical methods. It is possible to obtain an understanding of the " physical " ideas by the aid of comparatively simple mathematics. Furthermore, it is not necessary to follow the same methods of presentation of these ideas that were used

at the beginning by the pioneers in this field. This avoids the introduction of concepts which are both difficult to grasp and probably unessential, at least in the initial stages, for a comprehension of some of the basic principles and deductions. Professor E. T. Bell[1] has spoken of the "metaphors of quantum physics," and if we regard the mathematics of the new quantum theory as merely a symbolic language for the interpretation of physical phenomena, and not as a representation of the actual processes involved, then we shall find that a great deal of the apparent mystery disappears.

After all, it is essential to realize that quantum mechanics is merely the most convenient type of language which has been evolved, so far, for the representation of a large number of observations in physical science which have accumulated during the past three decades. It is a language in which there is a one-to-one correspondence between certain symbols and certain observations, and the mathematical technic constitutes the most logical method for deriving from these observations such conclusions as may be subjected to further experimental tests.

Consequently, it is essential that we should first consider carefully the actual observations which have led to the new point of view. The relationship that should exist between observations and their interpretation is one that has not always been clearly defined. It is comparatively easy to confuse the shadow with the substance, and what is often intended by the theoretical physicist as a working analogy is assumed by others to be an actual physical model of the new phenomena.

Some twenty years ago, H. Poincaré, one of the greatest mathematicians of that period, stated his opinions on this point in a work entitled "Science and Hypothesis." "Experiment," he wrote, "is the sole source of truth. It alone can teach us anything new; it alone can give us certainty. But to observe is not enough. . . . The scientist must set in order. Science is built up with facts, as a house is with stones. But a collection of facts is no more a science than a heap of stones is a house."

The scientist attempts to generalize from these observations, and thus sets up a theory so that he may be able to predict the results of new experiments. As has been stated in an address by I. Langmuir,[2]

> Our theories consist fundamentally in the setting up of a model which has properties analogous to the phenomena which we have observed. For example, Bohr, taking into consideration certain properties of hydrogen atoms, proposed a model for these atoms which consisted of an electron revolving in an orbit about a nucleus. The energy changes that would take place in this model, according to calculation, were found to be identical with those that were observed for hydrogen atoms. The theory was thus useful and served to explain the properties of hydrogen.

[1] E. T. Bell, *Sci. Monthly*, **32**, 193–209 (Mar., 1931).

[2] I. Langmuir, *Gen. Elec. Rev.*, **37**, 312 (1934).

Since any such model is an abstraction formed for a definite purpose, it is necessarily incomplete and therefore the model must never be confused with the physical phenomena which it represents. We should therefore never ask whether the model represents reality. It is sufficient to say that in certain respects it corresponds to reality. For example, although the equations which Bohr derived from a consideration of his model are still valid, we have today quite other explanations of the behavior of these atoms. The original Bohr model has lost its usefulness.

Even the atomic theory of matter, which is so universally accepted today, consists essentially in the setting up of a model, in which chemical compounds are conceived of as being made up of definite arrangements of atoms, to which we assign suitable properties. . . .

Most of the laws of physics are stated in mathematical terms. But a mathematical equation itself is nothing more than a kind of model. We establish, or assume, a correspondence between observable quantities and the symbols of an equation, and then, after a mathematical transformation, obtain a new relation or equation. If we can establish a similar correspondence between the symbols of the new equation and observational data obtained after an experiment has been performed, we have demonstrated the power of the mathematical theory to predict events. It thus becomes a useful theory.

Above all, it is necessary to realize that analogies and models are always limited in their scope, and conclusions, based on their use, must be tested constantly by further experiment.

In his book, " The Logic of Modern Physics," P. W. Bridgman has emphasized one guiding principle in the formation of the concepts for describing any new observations. " The concept," he states, " should be synonymous with the corresponding set of operations." He illustrates this statement by applying it to the physical concept length, and to the philosophical concept " absolute time," and then makes the following statement:

It is evident that if we adopt this point of view toward concepts, namely that the proper definition of a concept is not in terms of its properties but in terms of actual operations, we need run no danger of having to revise our attitude toward nature. For if experience is always described in terms of experience, there must always be a correspondence between experience and our description of it, and we need never be embarrassed, as we were in attempting to find in nature the prototype of Newton's absolute time. Furthermore, if we remember that the operations to which a physical concept are equivalent are actual physical operations, the concepts can be defined only in the range of actual experiment, and are undefined and meaningless in regions as yet untouched by experiment. It follows that strictly speaking we cannot make statements at all about regions as yet untouched, and that when we do make such statements, as we inevitably shall, we are making a conventionalized extrapolation, of the looseness of which we must be fully conscious, and the justification of which is in the experiment of the future.

Thus, in order to understand the function of our present theories on the structure and behavior of atomic and molecular systems and of electrons, it is essential to consider, first of all, the fundamental experi-

mental observations upon which these theories are based. Obviously it is possible, in such a discussion as the following, to mention only those facts which are both the most important and most readily understood. Furthermore, it is not at all necessary in presenting these observations to adhere to a historical sequence. It is much more essential to arrange these facts in order of increasing deviation from what might have been predicted on the basis of classical physics.

1.2 Energy States of Atomic Systems.[3] The atomic and molecular theories were found to be useful in the interpretation of chemical phenomena, and their utility was extended by the development of the kinetic theory of gases. Toward the end of the nineteenth century came the discovery of the electron, of X-rays, and of radioactive phenomena. While previous theories had led to the possibility of estimating the concentration and size of atoms or molecules, the new observations led to the conclusion that the atom itself is a complicated structure composed of electrons and positive charges. But it was not until about 1911 that a first really successful theory of atomic structure was suggested by Rutherford, and the subsequent investigation on X-ray spectra by Moseley, as well as those on isotopes by Aston, led to a new understanding of the periodic arrangement of the elements.

The model of an atom, consisting of a positively charged nucleus surrounded by one or more electrons, represented a significant departure from prevalent views in physics, since, on the basis of these views, such an atom must be inherently unstable. Nevertheless, it was only by means of this theory that the facts of radioactive disintegration and the observations on the scattering of alpha particles could be interpreted satisfactorily. The next problems to be investigated were manifestly those of electron configurations within the atoms themselves and of chemical combination between atoms.

The theory of the origin of spectral lines first suggested by N. Bohr in 1913 started new lines of investigations in the applications of quantum theory to atomic structure problems. One of the most striking of these early experiments was that carried out by J. Franck and G. Hertz in 1915. They showed that, when electrons are allowed to collide with atoms, there is a transfer of energy to the atoms only at certain critical values of the energy of the electron. If we designate the mass, charge, and velocity of the electron by μ, e, and v, respectively, the relation between the kinetic energy of the electron and the potential difference V through which it is accelerated, is given by

$$\tfrac{1}{2}\mu v^2 = Ve. \tag{1}$$

[3] See references to collateral reading at the end of the chapter.

Franck and Hertz found that this kinetic energy could be transferred completely to an atom only at certain critical values of V (critical potentials), thus indicating that for each type of atom there exists a certain series of discrete energy states. For collisions of atoms with electrons having an energy less than the lowest of these critical values (which we shall designate by V_r), the laws of elastic collisions apply, while for energies above the value V_r,[4] the electron loses only that amount of energy which corresponds to the next higher critical value.

A second observation, made by Eldridge in America and also by Franck and Hertz, was that an atom in the state corresponding to V_r is able to emit a monochromatic radiation of which the frequency ν is proportional to V_r in accordance with the relation

$$h\nu = V_r e, \tag{2}$$

where h is Planck's constant. It was also observed by these investigators and others that the frequency of *any* line, in the spectrum of an atomic system, could always be represented by a similar relation of the form

$$h\nu = (V_1 - V_2)e, \tag{3}$$

where V_1 and V_2 denote the energy values in volts corresponding to two different critical states of the atom as determined by bombardment with electrons.

Figure 1 illustrates these observations in the case of collisions between electrons and sodium atoms. As long as the energy of the electron is less than 2.10 volts, the collisions are perfectly elastic, and, in accordance with the laws of *ordinary mechanics* (applied to the collision between two particles), the electron loses only an insignificantly small fraction of its energy, namely, $2\mu/M$, where M = mass of atom. (In the case of sodium, $2\mu/M = 4.8 \times 10^{-5}$.) At 2.10 volts, or a slightly higher kinetic energy of the electron, an *inelastic collision* occurs.

The electron transfers a fraction of its energy, corresponding to 2.10 volts, to the sodium atom and under suitable conditions it will be observed that the vapor *emits the two D lines* of sodium of wave lengths 5890 and 5896. In other words, the 2.10 volts kinetic energy of the electron is used in exciting the sodium to the first excited state, which, as is shown in Fig. 1, is designated spectroscopically as 3P, and when the excited atom returns to the normal state, the radiation corresponding to the D lines is emitted. (Actually the 3P state consists of two states of slightly different energy contents — hence the emission of two lines of nearly the same wave length.)

[4] The energy value is, of course, $V_r e$; but it is customary to designate energy values in terms of electron volts, that is, the value of V as defined by equation (1). See Appendix II for values of universal constants and conversion factors.

As the electron energy is increased beyond 2.10 volts and maintained below 3.18 volts, any inelastic collision between an electron and a sodium atom results in the transfer to the latter of only 2.10 volts energy, while the *excess* over this value is retained by the electron as kinetic energy. At 3.18 volts, the electron can excite a sodium atom to the 4S state, and

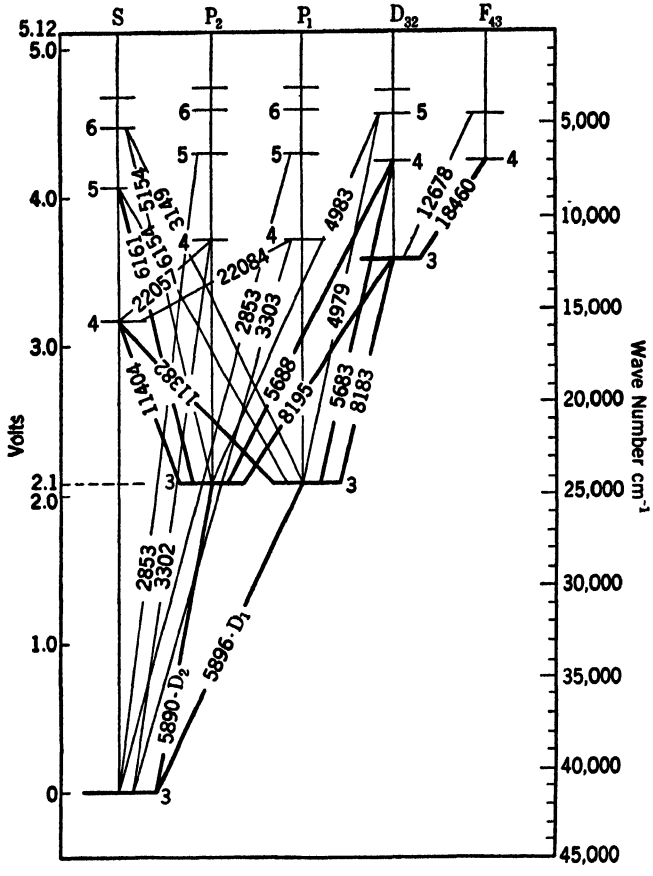


FIG. 1. Energy levels and lines in arc spectrum of sodium.

from this state the only transition which can occur is that to the 3P state, with the accompanying emission of the infra-red lines λ11,382 and λ11,404. (Transitions can occur only between S and P states, and P and D states, but not between states designated by similar letters.)

Thus, as the kinetic energy of the electron is increased, it becomes possible to excite the sodium atom to successively higher energy states (or "levels"), and the spectrum changes from a narrow doublet, observed when V is below 3.18 volts, to one consisting of an increasing

number of lines, until finally, when V becomes equal to or exceeds 5.12 volts, *all the lines in the arc spectrum of sodium* appear. Observation shows that, at this latter voltage, ionization occurs, with formation of Na^+ and an electron.

Even before the advent of the Bohr theory, spectroscopists had recognized that in some of the simpler spectra at least (such as those of H, the alkali metals, etc.), it is possible to represent the frequency ν of any line in the form

$$\nu = R\left(\frac{1}{a^2} - \frac{1}{m^2}\right), \tag{4}$$

where R is a universal constant (the Rydberg constant) and a is a constant for a given series of lines, while m increases by *integral values* with increase in ν for the different members of the series. In terms of the Bohr theory, the values R/a^2 and R/m^2 correspond to "energy levels," and in Fig. 1 a number of such energy levels are shown together with the spectral lines from which the energy levels have been deduced. Instead of designating the values of these levels in terms of the frequency, spectroscopists have used the *wave numbers* $\tilde{\nu} = \nu/c$, where c = velocity of light. Thus, the wave length of any line is given in terms of the wave numbers of the corresponding levels by the relation

$$\frac{1}{\lambda} = \tilde{\nu}_1 - \tilde{\nu}_2 \text{ (cm.}^{-1}\text{)}, \tag{5}$$

where $\tilde{\nu}_1$ is the wave number of the lower and $\tilde{\nu}_2$ that of the upper level, between which the transition occurs. In Fig. 1, the wave numbers of the different levels have been indicated, and it will also be observed that each column corresponds to energy levels belonging to the same spectral series. From the difference in wave numbers $\Delta\tilde{\nu}$ for any two levels, the corresponding electron volts may be derived by the relation

$$V = \frac{\Delta\tilde{\nu}}{8106}. \tag{6}$$

1.3 Resonance Radiation. While the first excited state of the sodium atom may be obtained by impact of a sodium atom with an electron having an energy greater than 2.10 volts, the same result may also be obtained by allowing the D lines from a sodium-vapor lamp to strike a heated bulb containing sodium vapor at low pressure and no electrodes whatever. The sodium atoms absorb the radiations λ5890 and λ5896, and thus become excited. On returning spontaneously to the normal state, the same wave lengths are reëmitted and the bulb containing the vapor glows with a faint yellow color. The analogy with resonance

phenomena in other fields of physics has led to the designation of this radiation as of the *resonance type*, and the value 2.10 volts required to excite sodium to the state at which it will emit this resonance radiation is therefore known as the *first resonance potential.*

By noting the voltages at which electrons lose kinetic energy by inelastic collisions with the atoms, and also by observing the wave lengths of the resonance radiation, it has been found possible to determine energy-level diagrams for most of the elements, not only in their normal states, but also at different stages of ionization. The importance, however, of all these observations from the point of view of the quantum theory is that they lead to two conclusions, which, as a matter of fact, were stated by Bohr in his original paper in the form of the following two postulates:

A. An atomic system can, and can only, exist permanently in a certain series of states corresponding to a discontinuous series of values for its energy, and consequently any change in the energy of the system, including emission and absorption of electromagnetic radiation, must take place by a complete transition between two such states. These states will be denoted as the "stationary states" of the system.

B. The radiation absorbed or emitted during a transition between two stationary states is monochromatic and possesses a frequency ν, given by the relation [analogous to (3)]

$$h\nu = E_n - E_m, \tag{7}$$

where E_n and E_m are the energies of the two different states.

These observations thus indicate that in the interaction of matter and radiation the magnitude of the energy interchange is measured in terms of a unit, or *quantum* as it has been designated, *which is proportional to the frequency* ν. That this involves an atomistic view of the nature of radiant energy is deduced from another series of investigations, *viz.*, those on the photoelectric and inverse photoelectric effect and on the Compton effect.

1.4 Photoelectric Effect. In the emission of electrons from metals by the incidence of radiation, it has been observed that the energy of the emitted electrons is proportional to the *frequency* of the radiation, and *not* to the intensity. If we let W denote the work required to pass the electron through the surface, then, according to Einstein, the maximum energy of the emitted electron is given by the relation

$$\tfrac{1}{2}\mu v^2 = h\nu - W, \tag{8a}$$

where ν is the frequency of the radiation used. The velocity of the electrons is measured, in general, by the magnitude of the retarding potential voltage V required to decrease the velocity to zero, in ac-

cordance with equation (1). Consequently, (8*a*) may be written in the more usual form

$$\begin{aligned} Ve &= h\nu - W \\ &= h\nu - h\nu_0, \end{aligned} \tag{8b}$$

where $\nu_0 = W/h$ is the minimum frequency which causes photoelectric emission from the given surface and is therefore known as the "photoelectric threshold."

Equation (8*b*) thus leads to the conception that the electron takes up a quantum of incident radiation, of magnitude $h\nu$, and uses it partly in passing through the surface, and partly in the form of acquired kinetic energy. Such a relation is inconsistent with any theory of spreading wave fronts of light. A single atom on the emitting surface is apparently able to concentrate the radiant energy incident on an area a million or more times greater than the atomic cross section, into a single unit (or quantum) and then utilizes this energy to eject an electron.

Similar observations have been made on the ionization of atoms by X-rays. Apparently the X-rays are capable of passing over billions of molecules without losing any energy, and then, by accident as it were, one molecule absorbs the energy of a whole train of X-ray waves with the resulting ejection of an electron (which constitutes the process of ionization). Furthermore, in this case also we find that the relation between frequency of X-ray radiation and velocity of emitted electrons is given by Einstein's relation, equation (8*a*).

The inverse photoelectric effect is another illustration of the application of the same relation. If a stream of electrons is directed against any solid, as in an X-ray tube, the maximum frequency of the radiation emitted increases linearly with the voltage in accordance with equation (8*b*).

These observations on the relation between radiant energy and kinetic energy of electrons are quite in disagreement with predictions based on the undulatory theory of light. "The effects," as Sir William Bragg has pointed out, "are as if the energy were conveyed from place to place in entities, such as Newton's old corpuscular theory of light provides." In other words, whereas the observations on interference and diffraction lead quite logically to a wave theory of light, the quantum phenomena discussed in the previous paragraphs can be interpreted only in terms of a corpuscular theory; that is to say that, in the interaction of radiation with electrons, the former behaves as if it were constituted of light units, or *photons*, as they have been designated. On this point of view we assume that these photons are guided by the electromagnetic waves and that what is distributed uniformly along the wave front is not the

energy but *rather the probability of occurrence of a photon.* For light waves of such intensity that the area covered by a million atoms is receiving one quantum of energy per unit time, there is a probability of one in a million that any one atom will be bombarded in that interval by a photon, with the resultant ejection of an electron.

1.5 The Compton Effect. This corpuscular conception of the nature of radiant energy was utilized by A. H. Compton in 1923 to interpret some very significant observations on the scattering of X-rays by solid bodies.

When X-rays impinge on matter, secondary rays of slightly longer wave length, that is, of lower frequency, are produced. This result is distinctly different from that observed for ordinary or visible light and could not be understood on the basis of any classical wave theory. However, A. H. Compton suggested an interpretation based on the corpuscular theory of light which has met with signal success.

If the incident X-rays of frequency ν be considered as a stream of light particles or photons, then each photon carries an amount of energy $E = h\nu$, and possesses a momentum P, which, in accordance with the observations on light pressure, is given by the relation $P = h\nu/c$, where c is the velocity of light. When a photon collides with a free or loosely bound electron, there occurs an interchange of both energy and momentum *in accordance with the laws of elastic collision for particles.* Consequently, the photon suffers a recoil in one direction with loss of momentum and decrease in energy, while the electron moves off in another direction with added momentum and increased kinetic energy. Such a collision is illustrated in Fig. 2, where θ is the angle between the direction of the incident and that of the scattered photon. Now the interesting point is that, while *the distribution of values of* θ *is governed by a law of probability,* the relation between decrease in frequency of the photon and the value of θ for any individual collision is that calculated on the basis of the laws of conservation of energy and momentum.

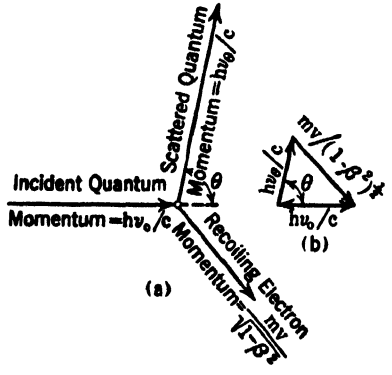


FIG. 2. Illustrating the theory of the Compton effect.

Here, then, we have a phenomenon which, like the photoelectric effect, can be explained only in terms of a corpuscular theory of energy. Yet, in all these observations, use is made of the wave theory to deter-

mine wave lengths and frequencies. We are thus led to adopt a *dualistic conception of the nature of radiant energy.* When dealing with interference, diffraction, and polarization phenomena, we find it necessary to use the undulatory theory; when dealing with the interaction of radiation and matter, it is necessary to use the corpuscular concepts, energy and momentum. The two apparently contradictory aspects are connected by the extremely significant relations

$$E = h\nu, \tag{9}$$

and

$$P = \frac{h\nu}{c} = \frac{h}{\lambda}. \tag{10}$$

1.6 Undulatory Phenomena Associated with Corpuscles. While de Broglie and Schroedinger had already suggested that matter also might partake in a similar dualistic behavior, the first experimental evidence for this idea was obtained by C. J. Davisson and L. H. Germer. In 1927 they made certain observations on the reflection of electrons from single crystals of nickel, which could be interpreted only on the assumption that under these conditions a stream of electrons possesses undulatory properties. The observations on the variation in intensity of the reflected beam with angle of incidence, for a homogeneous beam incident on the crystal, led to the conclusion that there exists, associated with the corpuscular kinetic energy of the electrons, a wave motion for which the wave length λ (known as the *de Broglie wave length*) is related to the momentum μv by an equation identical with that used by Compton, of the form

$$\lambda = \frac{h}{\mu v} = \frac{h}{\sqrt{2\mu V e}}, \tag{11}$$

where V is the potential difference through which the electrons are accelerated in acquiring the velocity v.

These observations were made with low-velocity electrons, but G. P. Thomson showed a little later that high-velocity electrons are diffracted by thin metal films in exactly the same manner as X-rays, thus repeating with electrons the famous experiment by which Laue had demonstrated in 1913 the wave nature of these rays. Subsequently it was shown by A. J. Dempster that, in the reflection of protons from crystal surfaces, the phenomena observed indicate, for this case also, a wave length associated with the corpuscular momentum which is given by equation (11).

Before discussing the significance of this relation, it is essential to consider what values of λ we may expect, on the basis of this equation, for

certain cases of corpuscular motion. Since $h = 6.56 \times 10^{-27}$ erg · sec., it follows that for $\mu = 1$ gm. and $v = 1$ cm. per sec., $\lambda = 6.56 \times 10^{-27}$ cm. This is much too small to be measured by any sort of grating available, and hence cannot be observed experimentally. From the investigations of crystal lattice structures by means of X-rays, it has been shown that the distances between atoms in such crystals are of the order of 10^{-8} cm. This therefore determines, for the magnitudes of corpuscular wave lengths that may be observed by means of crystals, values ranging from 10^{-7} to 10^{-10} cm., while optically ruled gratings enable us to measure wave lengths exceeding 10^{-6} cm. The values of μv corresponding to the lower range of wave lengths are $h/10^{-7}$ to $h/10^{-10}$, that is, from 6.55×10^{-20} to 6.55×10^{-17} gm. cm. sec.$^{-1}$, and such values are obtained only with atoms or electrons. For instance, in the case of a hydrogen molecule ($\mu = 3.3 \times 10^{-24}$ gm.), the velocity at room temperature is about 2×10^{5} cm. per sec., and therefore $\mu v = 6.6 \times 10^{-19}$ gm. cm. sec.$^{-1}$, while for an electron ($\mu = 9 \times 10^{-28}$ gm.) having a velocity 5.9×10^{8} cm. per sec. (corresponding to a fall through a potential of 100 volts), $\mu v = 5.3 \times 10^{-19}$ gm. cm. sec.$^{-1}$ and $\lambda = 1.24 \times 10^{-8}$ cm.

It is for these reasons that phenomena exhibiting the characteristics associated with waves may be observed experimentally only with such ultramicroscopic particles as atoms and electrons and cannot possibly be detected, at least in the light of present knowledge, with macroscopic corpuscles.

1.7 Principle of Indeterminacy. These effects which have been described rather briefly in the previous sections thus lead to conclusions which are apparently quite opposed to notions inherited from classical physics. Classical physics conceived light as an undulatory motion in a hypothetical ether; the theory of relativity discarded the ether, and it would appear that the quantum relations obliterate the waves. On the other hand, while the experiments on deflection of electrons in electrostatic and magnetic fields led physicists to assign to electrons a mass μ and a velocity v, as well as a charge e, the experiments on diffraction lead to the conception of "electron waves," to which a definite "wave length" may be assigned.

What explanation can be deduced for this seeming dualism in the behavior of both radiation and matter? The answer, first perceived most clearly by Heisenberg and Bohr, is that this dualism is actually inherent in the experimental arrangements used, in the agencies of observation themselves. *The nature of the experiment controls the result actually observed.*

The difficulty is this, that we have always assumed that we could

treat phenomena as something apart from the tools used in the observations. After all, as Eddington reminds us, " The world of physics is a world contemplated from within, surveyed by appliances which are part of it and subject to its laws," whereas it had been assumed that such observation revealed something that is independent of the mode of observation itself.

When we measure a length with a meter stick, or observe the position of an oil drop through a telescope, we are justified in assuming that the act of observing has introduced no effects on the object of observation. Consequently, it is possible, in ordinary dynamical problems, to specify the instantaneous state of a particle in terms of its position (which we shall designate by x) and its velocity v or, more accurately, its momentum $p = \mu v$. From a knowledge of the forces acting on the particle, it is then possible to predict its subsequent behavior, as, for instance, its position and velocity after any period of time t. Such a prediction is valid because it is possible to make observations on the initial conditions without " spoiling " the results of the measurements. " However," as Heisenberg has pointed out,[5]

> This assumption is not permissible in atomic physics; the interaction between observer and object causes uncontrollable and large changes in the system being observed, because of the discontinuous changes characteristic of atomic processes. The immediate consequence of this circumstance is that in general every experiment performed to determine some numerical quantity renders the knowledge of others illusory, since the uncontrollable perturbation of the observed system alters the values of previously determined quantities. If this perturbation be followed in its quantitative details, it appears that in many cases it is impossible to obtain exact determination of the simultaneous values of two variables, but rather that there is a lower limit to the accuracy with which they can be known.

For instance, in the Bohr theory of the hydrogen atom, the motion of an electron around the nucleus is treated on the same basis as the motion of the earth around the sun. It is assumed that we can measure both the position and velocity of the electron at any instant and that from this we can derive a magnitude which we designate as frequency of revolution in an orbit. But can position and velocity be specified simultaneously for an electron in an atomic system? Heisenberg's answer is that this is impossible. In fact, the more accurately we attempt to determine the position, the less accuracy we attain in the measurement of velocity, and vice versa.

As an illustration, let us consider the manner in which we might try to determine the position at any instant of an electron in motion. In order to see the electron, it must be illuminated, and from optical con-

[5] W. Heisenberg, "The Physical Principles of the Quantum Theory," p. 3.

siderations it is known that for an ideal lens the uncertainty Δx in the determination of x is given by the relation

$$\Delta x = \frac{\lambda}{\sin \theta}, \tag{12}$$

where λ is the wave length of light used and 2θ is the aperture of the lens (see Fig. 3). Thus θ should be chosen as large as possible, and λ as small as possible. Theoretically we could use gamma rays, the shortest wave lengths of radiation obtainable. To make the observation it is necessary that at least one photon should be scattered by the electron and pass through the microscope lens to the eye of the observer. In consequence of the Compton effect, the electron receives a recoil, and the amount of this recoil cannot be determined since the lens receives in the same focus all the rays originating in the angle 2θ. Thus the uncertainty in the magnitude of the loss in momentum of the electron in the x-direction is given by

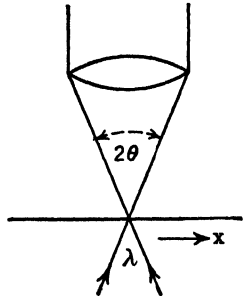


FIG. 3. Illustrating the Principle of Indeterminacy.

$$\Delta p_x \geqslant \frac{h}{\lambda} \sin \theta. \tag{13}$$

The inequality indicates that the magnitude of the inexactitude will never be less than $h \sin \theta/\lambda$, but may be greater, owing to physical imperfections in the experimental arrangement.

Consequently, we arrive at the very significant result

$$\Delta x \cdot \Delta p_x \geqslant h. \tag{14}$$

To minimize the loss in momentum, we might use radiation of much greater wave length. In fact, we might attempt to measure the velocity of the electron by means of the Doppler effect, and in order to increase the accuracy of observation, it would be necessary to work with very long-wave-length radiation, but this would increase the inaccuracy in determination of position.

In the foregoing discussion, use has been made, on the one hand, of the wave theory in connection with resolving power of the lens, and, on the other hand, of the corpuscular theory of the Compton effect. However, the conclusion stated in equation (14) may be derived also from a consideration of the wave properties exhibited by electrons when made to pass through a slit. We imagine a homogeneous beam of electrons incident in the normal direction on a screen containing a slit.

In order to fix the position of one of the electrons at the instant of passing, we must choose a slit of extremely narrow width d. The coördinates parallel to the screen are thus determined with the accuracy

$$\Delta x = d.$$

But if d is comparable in magnitude with the de Broglie wave length λ, the electrons will be deflected at the edges (diffraction phenomenon). Consequently, the emergent beam has a finite angle of divergence θ, which, according to the laws of optics, is determined by the relation

$$\sin \theta = \frac{\lambda}{d} = \frac{\lambda}{\Delta x},$$

and the momentum of the electrons in a direction parallel to the screen is uncertain, after passing the slit, by an amount

$$\Delta p = \frac{h}{\lambda} \sin \theta.$$

From these two relations, equation (14) follows, as before.

A similar relation is valid for the simultaneous determination of E, the energy, and t, the instant of observation. For, in order to determine the difference in frequency $\Delta\nu$ between two frequencies ν and $\nu + \Delta\nu$, we must extend the observation over an interval of time $\Delta t = 1/\Delta\nu$. Hence

$$h\Delta\nu\Delta t = \Delta E \cdot \Delta t \geqslant h. \tag{15}$$

The conclusions stated in equations (14) and (15) constitute the generalization which is known as Heisenberg's *Principle of Indeterminacy*,[6] and though it does not enable us to make any calculations on the behavior of atomic systems and electrons, it is extremely important in indicating the nature of the predictions which can be made about such particles.

Heisenberg's principle postulates that there exists a *fundamental limitation* governing the possibility of associating exact determination of position with exact determination of momentum, when dealing with such systems as atoms and electrons, and the reason for this is the fact that any observation on atomic systems or electrons involves an interaction with agencies of observation, not belonging to the system. Thus the *initial conditions* in any dynamical problem involving atoms are *indeterminable* to the extent defined by equation (14), and consequently we cannot expect classical methods to be valid for calculating the behavior of a microscopic system such as an atom or an electron.

[6] What Heisenberg designated as "Unbestimmtheit" has been translated into the English equivalents: inexactitude, indefiniteness, uncertainty, and indeterminacy.

That this limitation is of negligible significance in the calculation of macroscopic systems is readily evident from considerations similar to those advanced previously in the discussion of the undulatory phenomena associated with corpuscular motion. The experimental limitations in the accurate determination of either position or velocity, in dealing with the motion of ordinary masses, are so large that the Heisenberg inexactitude relation becomes completely obscured. This is no longer true, however, when dealing with the motions of electrons in atomic systems. In view of the impossibility of determining accurately the initial conditions in these cases, a precise statement of subsequent occurrences is no longer possible. What, then, can be calculated with regard to the behavior of such a system?

In the ordinary affairs of life we have learned to solve such problems by applying the methods of the theory of probability. Thus the life of any individual human being is indeterminate in duration, but life insurance statistics enable us to state the life expectancy for any individual at a given age. Similarly, in the manufacture of any piece of mechanism, where such production involves a large number of units, it is possible to predict on the basis of statistical information what the probability is for the occurrence in any unit of a given type of characteristic.

In the kinetic theory of gases we have the well-known probability distribution laws of Maxwell and Boltzmann. These laws state the manner in which the probability of occurrence of a given range of velocities or energies varies with the velocity or energy. Thus it is found that, while there is a decreasing probability for the occurrence of very high or very low velocities, there exists, for each temperature and composition of gas, a certain velocity for which the probability is a maximum.

Now let us return to the consideration of the problem in atomic mechanics. Here, as has been mentioned previously, we are confronted with the fact that initial conditions are defined only within the limits determined by Heisenberg's principle. In view of this uncertainty, it is evident that all that we may expect to determine from the solution of a problem on the behavior of an atomic system is the probability of occurrence of any individual event. That is, the new quantum mechanics is essentially *a technic for the calculation of statistical probabilities*, and not one which enables us to predict absolute certainties in the same sense as we have been led to expect from ordinary mechanics.

Since, as has been emphasized previously, the indeterminacy becomes less and less significant with increase in the values of μv and x beyond those dealt with in the consideration of atomic systems, it is evident that for *macroscopic* phenomena the new quantum mechanics must yield results which are identical with those derived by classical mechanics.

Bohr had recognized the necessity for fulfilling this condition in developing his hybrid theory of classical mechanics and quantizing principles. As a result he formulated his famous Correspondence Principle, and, in the new quantum theory as well, the spirit of this principle has been maintained. For instance, as shown by both C. G. Darwin[7] and E. H. Kennard,[8] the path of a macroscopic particle, falling freely under the action of gravitational forces, when derived by the methods of Schroedinger, is found to be identical in form with that derived by the method of Newtonian mechanics. If we adopt the language of the quantum theory, we may describe this result thus: there is an infinitely high degree of probability that, at the end of a given interval of time, the magnitudes of μv and x will approach certain determined values more closely than any differences that can be measured, even with the utmost possible physical precision. On the other hand, the ordinary calculation states that, at the end of the given interval of time, μv and x will have these actual definite values. In other words, for large-scale phenomena, classical mechanics merely states as a certainty a result to which quantum mechanics assigns such *an extremely high degree of probability that for all practical purposes it becomes a certainty.*

COLLATERAL READING

(For exact titles and names of publishers see Appendix I.)

The reader should supplement the rather brief review of quantum phenomena given in the preceding chapter by consulting the following books and articles:

1. Ruark, A. E., and Urey, H. C., "Atoms, Molecules and Quanta."
2. Richtmyer, F. K., "Introduction to Modern Physics."
3. Darrow, K. K., "Contemporary Physics."
4. Darwin, C. G., "The New Conceptions of Matter."
5. Dushman, S., Chapter XVI of Taylor's "Treatise on Physical Chemistry."

An interesting review of the observations which led to the formulation of quantum mechanics is given by Lindsay and Margenau, "Foundations of Physics," Chapter IX, Section 1. R. W. Gurney, "Elementary Quantum Mechanics," Chapter IV, discusses the principle of indeterminism, and N. Bohr, supplement to *Nature*, April 14, 1928, p. 580, has given a most stimulating presentation of this principle and its physical significance. S. Dushman, *Gen. Elec. Rev.*, **33**, 328, 394 (1930), has given a review of quantum phenomena and discussed the physical interpretation of the concepts of the new quantum mechanics, in non-mathematical language.

For a discussion of the *Compton effect*, the following original publications should be consulted:

1. Compton, A. H., *Phys. Rev.*, **21**, 207, 483 (1923); *ibid.*, **22**, 409 (1923); *J. Optical Soc. Am.*, **16**, 71 (1928); *Phys. Rev. Suppl.*, **1**, 74 (1929).

[7] C. G. Darwin, *Proc. Roy. Soc.*, **A117**, 258 (1928).
[8] E. H. Kennard, *J. Franklin Inst.*, **207**, 47 (1929).

This topic is also presented in an interesting manner by K. K. Darrow, *Bell System Tech. J.*, **8**, 64 (1928).

For a description of the experiments on *electron diffraction* the following publications are important:

1. DAVISSON, C. J., and GERMER, L. H., *Phys. Rev.*, **30**, 705 (1927); DAVISSON, C. J., *Bell System Tech. J.*, **7**, 90 (1928) and *J. Franklin Inst.*, **205**, 597 (1928).
2. THOMSON, G. P., *Proc. Roy. Soc.*, **A117**, 600 (1928); *ibid.*, **A119**, 651 (1928); *Science*, **70**, 541 (1929); *Nature*, **122**, 279 (1928).

CHAPTER II

THE SCHROEDINGER EQUATION IN ONE DIMENSION

2.1 The Concepts of Quantum Mechanics. It is the concept of probability that characterizes the new point of view, and quantum mechanics represents a modification of classical mechanics which enables us to predict the behavior of any system from the point of view of the Principle of Indeterminacy. As a consequence of the application of this new mechanics, certain conclusions have been deduced which are quite different from those expected on the basis of Newtonian mechanics. On the other hand, the new theory leads to discrete energy states for atomic systems by a much less artificial mode of derivation than was possible in the Bohr-Sommerfeld theory. Furthermore, the calculation now leads to correct solutions in those cases where the latter theory failed to give a satisfactory answer, while it yields the same results as the older theory wherever the latter did give results in agreement with observation.

The actual mathematical technic of the new quantum mechanics has evolved from what appeared at first to be three different lines of attack. The first one, originated by W. Heisenberg, is a purely symbolic type of mathematics and is quite unsuitable for elementary presentation. The second one, developed by E. Schroedinger (we shall use the symbol S. in referring to him), has enjoyed greater popularity, probably, as Eddington has suggested, " because it is the only one that is simple enough to be misunderstood." The third line of development is that presented by P. Dirac in his treatise, " The Principles of Quantum Mechanics "; one which involves, like Heisenberg's treatment, a symbolism which is apt to repel any who are not " mathematically minded."

In the presentation which the writer has attempted in the following sections, only the most essential aspects of the S. technic are considered, without regard to the particular arguments by which S. actually developed his equation. The reader who desires to obtain an idea of the actual method of derivation used by S. will find this presented in his original papers and some of the treatises on quantum mechanics.[1]

However, since it is impossible to present even the simplest formulation of the S. theory without recourse to certain fundamental mathe-

[1] See the list of "General References on Quantum Mechanics," in Appendix I; also E. Schroedinger, *Ann. Physik*, [4] **79**, 361, 489 (1926).

matical ideas, it has been considered advisable to precede the derivation of the S. equation and the consideration of some of its applications by some remarks on these more purely mathematical aspects.

2.2 Some Fundamental Differential Equations.[2] A differential equation is a quantitative expression of a hypothesis regarding the mechanism of a given phenomenon. In general, what we measure as " pointer-readings " are the results over a period of time, or over a finite distance, of certain forces acting between the bodies constituting the system under observation. These forces vary with time and with changes in relative arrangement of the parts of the system, but from the integration of the differential equation, whenever this is feasible, it is possible to deduce the total change in the magnitudes defining the state of the system under any given conditions or at any instant of time in terms of the initial conditions.

(1) Let us consider the very simple law governing the motion of a body under the action of a constant force. Let F denote the force, μ the mass, and s the distance measured along the path of motion.

By definition, the force is equal to the rate of change of momentum. Hence

$$F = \frac{d(\mu v)}{dt}, \tag{1}$$

where v = velocity at any instant.

For velocities which are small compared with that of light, μ is constant. Therefore, we can write

$$F = \mu \cdot \frac{dv}{dt}.$$

But velocity is defined as rate of change of distance along the path. That is,

$$v = \frac{ds}{dt}.$$

Consequently, we obtain the result

$$F = \mu \cdot \frac{d}{dt}\left(\frac{ds}{dt}\right) = \mu \cdot \frac{d^2s}{dt^2}. \tag{2}$$

Since F is constant, we obtain *the differential equation of the second order,*

$$\frac{d^2s}{dt^2} - \frac{F}{\mu} = 0 \quad \text{or} \quad \frac{d^2s}{dt^2} - \alpha = 0, \tag{3}$$

where α is the " acceleration."

[2] See list of references on mathematics, at the end of the chapter.

We now proceed to find s as a function of t. This will give us a "solution" of the differential equation. Integrating once, we have

$$\int_0^t d\left(\frac{ds}{dt}\right) = \alpha \int_0^t dt,$$

so that

$$\left(\frac{ds}{dt}\right)_t - \left(\frac{ds}{dt}\right)_0 = \alpha t$$

or

$$\frac{ds}{dt} = \alpha t + v_0,$$

where $v_0 = (ds/dt)_0$ is the velocity at $t = 0$. This is a *differential equation of the first order*, and integration of this equation gives the result

$$s = \tfrac{1}{2} \cdot \alpha t^2 + v_0 t + s_0, \tag{4}$$

where s_0 is the value of s at $t = 0$.

Figure 4 shows plots of (1) $F/\mu = d^2s/dt^2 = \alpha$, the acceleration; (2) $v = ds/dt$, the velocity; and (3) s, the distance traversed in time t, all as functions of t.

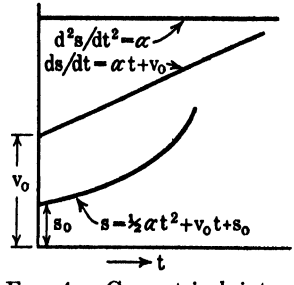


FIG. 4. Geometrical interpretation of a simple differential equation and of its solution.

Equation (4) is designated a *general solution* of equation (2), and it will be observed that, in the process of solving the latter, *two constants of integration* are introduced, one of which corresponds to the initial value of ds/dt and the other to the initial value of s. The differential equation (3) expresses the equation of motion of a body under the action of gravity, where $F/\mu = g$, the gravitational constant.

(2) The special case of equation (3) in which F, and consequently, α, is equal to zero, gives rise to the second-order differential equation of the form

$$\frac{d^2y}{dx^2} = 0, \tag{5}$$

which leads to the first-order equation

$$\frac{dy}{dx} = c, \text{ a constant,}$$

and this in turn, to the result

$$y = cx + d,$$

where d is a second integration constant.

The last equation is that of a straight line, and c defines the slope, while d defines the intercept on the y-axis, since $y = d$ for $x = 0$.

(3) A very important type of second-order differential equation is that representing a *simple harmonic motion.* If a particle of mass μ moves with respect to a fixed point in such a manner that the restoring force acting on the particle is proportional to the displacement (r) from the fixed point, the differential equation for the motion has the form

$$F = \mu \frac{d^2r}{dt^2} = -kr, \tag{6}$$

where k is a constant which we shall regard as positive.

In order to integrate this equation, we try the solution

$$r = r_0 \epsilon^{mt}.$$

Hence
$$\frac{d^2r}{dt^2} = r_0 m^2 \epsilon^{mt},$$

and substituting in (6) we obtain the " characteristic equation "

$$\mu m^2 + k = 0,$$

that is,
$$m^2 + \frac{k}{\mu} = 0.$$

The roots of this equation are $m = \pm i\sqrt{k/\mu} = \pm i\omega$, where $\omega = \sqrt{k/\mu}$, and $i = \sqrt{-1}$.

Consequently, the *general solution* or *complete integral* has the form

$$r = C\epsilon^{i\omega t} + D\epsilon^{-i\omega t}, \tag{7}$$

where C and D are two arbitrary constants of integration.

The right-hand side of equation (7) may be expressed in a more familiar form by making use of the series expansion formulas[3] for $\epsilon^{i\omega t}$ and $\epsilon^{-i\omega t}$

$$\epsilon^{i\omega t} = 1 + i\omega t + \frac{(i\omega t)^2}{1 \cdot 2} + \frac{(i\omega t)^3}{1 \cdot 2 \cdot 3} + \frac{(i\omega t)^4}{1 \cdot 2 \cdot 3 \cdot 4} + \cdots$$

$$\epsilon^{-i\omega t} = 1 - i\omega t + \frac{(i\omega t)^2}{1 \cdot 2} - \frac{(i\omega t)^3}{1 \cdot 2 \cdot 3} + \frac{(i\omega t)^4}{1 \cdot 2 \cdot 3 \cdot 4} - \cdots$$

[3] These expansions are derived from Maclaurin's series, which in turn is a special application of Taylor's theorem.

Therefore,

$$\epsilon^{i\omega t} + \epsilon^{-i\omega t} = 2\left\{1 + \frac{(i\omega t)^2}{1\cdot 2} + \frac{(i\omega t)^4}{1\cdot 2\cdot 3\cdot 4} + \cdots\right\}$$

$$= 2\left\{1 - \frac{(\omega t)^2}{1\cdot 2} + \frac{(\omega t)^4}{1\cdot 2\cdot 3\cdot 4} - \cdots\right\}$$

$$= 2\cos \omega t, \tag{8}$$

and

$$\epsilon^{i\omega t} - \epsilon^{-i\omega t} = 2i\left\{\omega t - \frac{(\omega t)^3}{1\cdot 2\cdot 3} + \frac{(\omega t)^5}{1\cdot 2\cdot 3\cdot 4\cdot 5} - \cdots\right\}$$

$$= 2i \sin \omega t. \tag{9}$$

That is,

$$\epsilon^{i\omega t} = \cos \omega t + i \sin \omega t \tag{10}$$

$$\epsilon^{-i\omega t} = \cos \omega t - i \sin \omega t. \tag{11}$$

In consequence, equation (7) may be expressed in the form

$$r = A \cos \omega t + B \sin \omega t$$

$$= A \cos \sqrt{\frac{k}{\mu}} \cdot t + B \sin \sqrt{\frac{k}{\mu}} \cdot t, \tag{12}$$

where the constants A and B are given in terms of C and D by the two sets of relations

$$A = C + D; \qquad B = i(C - D) \tag{13}$$

$$2C = A - iB; \quad 2D = A + iB. \tag{14}$$

That is, in the general case, C and D are *complex conjugate* quantities. If, as is customary, we designate the complex conjugate of any quantity by a bar over the symbol (or an asterisk), then it follows from equation (14) that

$$\bar{C} = D, \quad \text{and} \quad \bar{D} = C,$$

while

$$C\bar{C} = D\bar{D} = \frac{A^2 + B^2}{4}. \tag{15}$$

The quantity on the right-hand side of the last equation is *always real* and is designated the *norm* of the complex quantity C or D. The positive value of the square root, that is $\sqrt{A^2 + B^2}/2$, is known as the *modulus* of C or D. This is usually expressed in the form

$$|C| = |D| = \frac{\sqrt{A^2 + B^2}}{2} \tag{16}$$

We shall now consider the physical interpretation of equations (7) and (12). Evidently we can write the latter in the form

$$r = \sqrt{A^2 + B^2}\left\{\frac{A}{\sqrt{A^2 + B^2}} \cos \omega t + \frac{B}{\sqrt{A^2 + B^2}} \sin \omega t\right\}$$

$$= \sqrt{A^2 + B^2} \sin (\omega t + \delta), \tag{17}$$

where $\sin \delta = A/\sqrt{A^2 + B^2}$ and $\cos \delta = B/\sqrt{A^2 + B^2}$.

As t is varied from $t = 0$ to $t = 2\pi/\omega$, r passes through the cycle of values indicated in the following table.

t	r
$-\frac{\delta}{\omega}$	0
0	$\sqrt{A^2 + B^2} \cdot \sin \delta = A$
$\frac{1}{\omega}\left(\frac{\pi}{2} - \delta\right)$	$\sqrt{A^2 + B^2} \cdot \sin \frac{\pi}{2} = \sqrt{A^2 + B^2}$
$\frac{\pi}{2\omega}$	$\sqrt{A^2 + B^2} \cdot \cos \delta = B$
$\frac{1}{\omega}\left(\pi - \delta\right)$	0
$\frac{1}{\omega}\left(\frac{3\pi}{2} - \delta\right)$	$-\sqrt{A^2 + B^2}$
$\frac{2\pi}{\omega}$	A

In the *period* $\tau = 2\pi/\omega$, the value of r varies from 0 to $\pm\sqrt{A^2 + B^2}$, and it is evident that, for any value of $t = n\tau$ (where $n = 0, 1, 2,$ etc.), the value of r will pass through such a series of oscillations n times. That is, the particle performs harmonic vibrations about the point $r = 0$, of which the *amplitude* is $\sqrt{A^2 + B^2}$, and the frequency

$$\nu = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}. \tag{18}$$

The angle δ is known as the *phase angle*, and the particular value of this angle depends upon the initial conditions. In the case of equation (17), it is evident that the initial condition was $r = A$ for $t = 0$, and

this corresponded to the value

$$\delta = \tan^{-1} \frac{A}{B}.$$

On the other hand, for the initial condition $r = B$, the value of the phase angle would be

$$\delta = \tan^{-1} \frac{B}{A}.$$

Equation (17) expresses in the form of a single harmonic function the same motion as that expressed by the sum of two harmonic functions in (12). That the equations are equivalent may be demonstrated readily by considering the case $A = B$. Then the two equations become

$$x = A(\cos \omega t + \sin \omega t), \tag{i}$$

and, since $\sin \pi/4 = \cos \pi/4 = 1/\sqrt{2}$,

$$x = A\sqrt{2} \sin \left(\omega t + \frac{\pi}{4}\right). \tag{ii}$$

In the first of these equations, x is expressed as the sum of two identical harmonic waves which are 90° out of phase; in the second equation, x is expressed as a single harmonic wave, having an amplitude $\sqrt{2}$ times that of each of the waves in (i) and 45° out of phase with each of these.

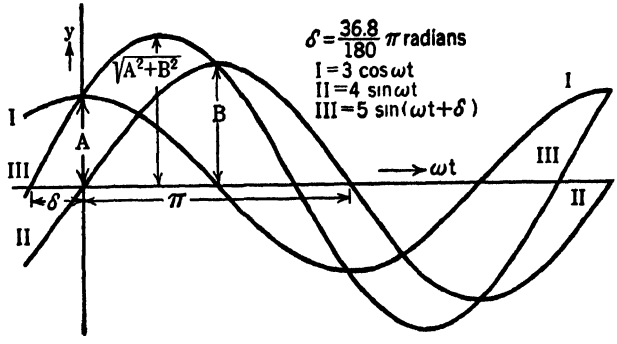


FIG. 5. Illustrating the superposition principle of wave motions.

The significance of these considerations is illustrated by the plots in Fig. 5. Curve I corresponds to $y = 3 \cos \omega t$; curve II to $y = 4 \sin \omega t$, and curve III to $y = 5 \sin (\omega t + \delta)$ where $\delta = 36.8°$. It will be observed that each ordinate for curve III is the sum of the ordinates, for the same value of t, for curves I and II. This illustrates a generalization

known as the *Principle of Superposition*, which is of extremely great importance in quantum mechanics.

We could also incorporate the phase angle δ into the exponential expressions used in equation (7). In that case, we put

$$r = C_0\epsilon^{i(\omega t+\delta)} + D_0\epsilon^{-i(\omega t+\delta)},$$

where $$2C_0 = (A - iB)\epsilon^{-i\delta},$$

and $$2D_0 = (A + iB)\epsilon^{i\delta}.$$

This leads to the result

$$A^2 + B^2 = 4\bar{C}_0C_0$$

and $$\sqrt{A^2 + B^2} = 2 \mid C_0 \mid = 2 \mid C \mid,$$

as before.

Returning to the consideration of equation (12), it follows that the velocity of the particle at any instant is given by

$$\frac{dr}{dt} = \dot{r} = -A\omega \sin \omega t + B\omega \cos \omega t,$$

and the initial momentum $p_0 = \mu\dot{r}_0$ is given by

$$p_0 = B\mu\omega.$$

Consequently, we can write equation (12) in the form

$$r = r_0 \cos \omega t + \frac{p_0}{\mu\omega} \sin \omega t, \tag{19}$$

and the momentum at any instant is given by

$$p = \mu\dot{r} = -r_0\omega\mu \sin \omega t + p_0 \cos \omega t. \tag{20}$$

Again, it is often convenient to express the motion in terms of the total energy E and the frequency.

From equation (6) it follows that the potential energy

$$V(r) = -\int_0^r F dr = \frac{kr^2}{2},$$

since by definition the potential energy[4] is the increase in energy due to the displacement of the particle from $r = 0$ to r.

Also, the kinetic energy at the point r is given by

$$T = \tfrac{1}{2}\mu\dot{r}^2 = \frac{p^2}{2\mu}.$$

[4] See more comprehensive discussion of potential and kinetic energy in Chapter IV.

Hence,

$$E = T + V = \frac{p^2}{2\mu} + \frac{kr^2}{2}, \tag{21}$$

and since this must be a constant for the system, it follows that p *as a function of* r *will be represented by an ellipse.*

Comparing equation (21) with the equation of an ellipse in terms of the major and minor semi-axes a and b, which is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{22}$$

it is seen that the semi-axes of the ellipse will be given in the present case by $\sqrt{2\mu E}$ for the axis along which p is measured, and by $\sqrt{2E/k}$ for the axis along which r is measured.

Since $k = \mu\omega^2$, it follows from equation (21) that

$$p^2 + \mu^2\omega^2 r^2 = 2\mu E, \tag{23}$$

which expresses p and r in terms of E.

By substituting p_0 and r_0 in this equation it follows that equation (20) may be written in the form

$$p = \sqrt{2\mu E}\cos(\omega t + \delta), \tag{24}$$

where $$\cos\delta = \frac{p_0}{\sqrt{2\mu E}}$$

and $$\sin\delta = \frac{\mu\omega r_0}{\sqrt{2\mu E}}.$$

Also it follows from (24) and (21) that we can write (19) in the form

$$r = \sqrt{\frac{2E}{k}}\sin(\omega t + \delta). \tag{25}$$

These two equations thus express p and r in terms of the total energy and of $\omega = 2\pi\nu$.

(4) Lastly, we shall consider the second-order differential equation of the form

$$\frac{d^2y}{dx^2} - m^2 y = 0, \tag{26}$$

where m is a constant, and m^2 is always positive.

As in the case of equation (6) we try the solution

$$y = A\epsilon^{kx},$$

so that[5]

$$\frac{d^2y}{dx^2} = y'' = k^2y.$$

Hence, the " characteristic equation " is

$$k^2 - m^2 = 0,$$

which has the roots $k = \pm m$, and this leads to the general solution for equation (26) of the form

$$y = A\epsilon^{mx} + B\epsilon^{-mx}. \tag{27}$$

For $x = 0$, $y_0 = A + B$, and

$$y_0' = \left(\frac{dy}{dx}\right)_{x=0} = m(A - B).$$

Thus the two integration constants can be expressed in terms of the values of y and y' for $x = 0$.

Corresponding to the exponential expressions for the sine and cosine functions given in (8) and (9), we have the following relations:

$$\epsilon^{mx} + \epsilon^{-mx} = 2\cosh mx, \tag{28}$$

$$\epsilon^{mx} - \epsilon^{-mx} = 2\sinh mx, \tag{29}$$

where cosh and sinh denote the " hyperbolic " cosine and " hyperbolic " sine functions, respectively. Hence (27) may be written in the form, analogous to (12),

$$y = (A + B)\cosh mx + (A - B)\sinh mx, \tag{30}$$

where $(A + B)$ and $(A - B)$ evidently may be replaced by two new constants, thus retaining the general form of solution with two arbitrary constants.

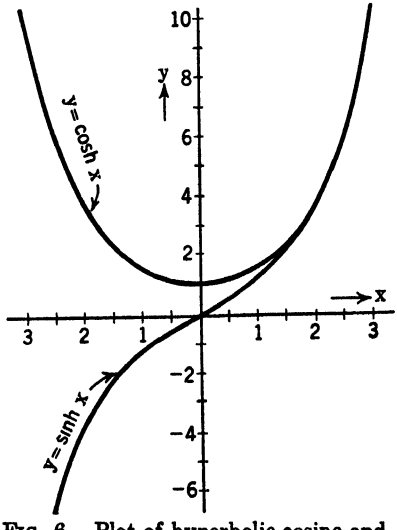


FIG. 6. Plot of hyperbolic cosine and sine functions.

Figure 6 shows the graphs for cosh x and sinh x, and it will be observed that the two functions become more and more nearly equal to each

[5] The notation $y' = dy/dx$, $y'' = d^2y/dx^2$; $\dot{y} = dy/dt$, $\ddot{y} = d^2y/dt^2$, is used very frequently in mathematical treatises.

other for large positive values of x, while they diverge more and more with increasingly negative values of x.

2.3 Equations for the Propagation of Wave Motion.[6] We shall now consider the manner in which it is possible to express the propagation of a wave motion along a single coördinate axis. Such an expression should give the amplitude of the wave at any point x as a function of the time t.

Let us consider the expression

$$y = y_0 \cos (\alpha x - \omega t), \tag{31}$$

where α and ω are two (positive) constants.

For values of $t = \alpha x/\omega$, the amplitude y repeats itself. Hence, equation (31) must represent the propagation of a wave motion for which the velocity of propagation (phase velocity) is $u = x/t = \omega/\alpha$.

If ν denote the frequency, and λ the wave length,

$$u = \nu\lambda,$$

and hence

$$\frac{\omega}{\alpha} = \frac{2\pi\nu}{\alpha} = \nu\lambda,$$

so that

$$\alpha = \frac{2\pi}{\lambda} = 2\pi\sigma, \tag{32}$$

where

$$\sigma = \frac{1}{\lambda} = \text{wave number.}$$

Consequently we can write (31) in the more convenient form

$$y = y_0 \cos 2\pi(\sigma x - \nu t). \tag{33}$$

If x becomes more positive, and t is made more positive to such an extent that $t = \sigma x/\nu$, the amplitude y repeats itself. Therefore, equation (33) must represent a wave motion for which the *direction of propagation* is the same as that of increase in the value of x, that is, *from left to right.*

On the other hand, if we wish to indicate propagation from *right to left,* it is evident that the corresponding expression must be

$$y = y_0 \cos 2\pi(\sigma x + \nu t), \tag{34}$$

since, in this case, for positive increase in t, x must be made negative in order to make y repeat itself.

[6] See references at end of chapter.

It is also evident that for propagation from *left to right* we could use the relation

$$y = y_0 \cos 2\pi(\nu t - \sigma x), \tag{35}$$

since $\cos\theta = \cos(-\theta)$, while for propagation from *right to left*, we could use the relation

$$y = y_0 \cos(-2\pi\nu t - 2\pi\sigma x). \tag{36}$$

If now we consider the expressions

$$\begin{aligned} y &= y_0 \cos 2\pi(\sigma x - \nu t) + i \sin 2\pi(\sigma x - \nu t) \\ &= y_0 \epsilon^{2\pi i\sigma x} \epsilon^{-2\pi i\nu t} \end{aligned} \tag{37}$$

and

$$y = y_0 \epsilon^{-2\pi i\sigma x} \epsilon^{2\pi i\nu t}, \tag{38}$$

we observe that we could use *either* relation to express a propagation from left to right, while for motion in the opposite direction we could use either of the two following relations:

$$y = y_0 \epsilon^{2\pi i\sigma x} \epsilon^{2\pi i\nu t} \tag{39}$$

or

$$y = y_0 \epsilon^{-2\pi i\sigma x} \epsilon^{-2\pi i\nu t}. \tag{40}$$

Now in quantum mechanics it is conventional, and there are also definite reasons for the procedure, as will be pointed out subsequently, for choosing those expressions in which $\epsilon^{-2\pi i\nu t}$ occurs. Under these conditions we shall regard $\epsilon^{2\pi i\sigma x}$ as indicating a wave motion from *left to right*, as follows from equation (37), while we shall regard $\epsilon^{-2\pi i\sigma x}$ as indicating propagation from *right to left*, as follows from equation (40).

We may also proceed now to derive from these relations the *differential equation for the propagation of a wave motion.*

Consider the relation

$$y = y_0 \epsilon^{2\pi i(\sigma x - \nu t)}.$$

Then, we derive the following partial differential coefficients:

$$\frac{\partial y}{\partial x} = 2\pi i\sigma y; \quad \frac{\partial^2 y}{\partial x^2} = -4\pi^2\sigma^2 y.$$

The symbol ∂ is used to designate differentiation of a function of more than one independent variable, with respect to one of these variables, maintaining the other variable constant. In the present case $y = y(x, t)$, and hence arises the necessity for using the notation for partial differentiation.

Similarly, we derive the relations

$$\frac{\partial y}{\partial t} = -2\pi i\nu y; \quad \frac{\partial^2 y}{\partial t^2} = -4\pi^2\nu^2 y.$$

Hence,

$$\frac{\partial^2 y}{\partial x^2} = \frac{\sigma^2}{\nu^2} \cdot \frac{\partial^2 y}{\partial t^2}$$

$$= \frac{1}{u^2} \cdot \frac{\partial^2 y}{\partial t^2},$$

where $u = \nu\lambda = \nu\sigma$ = phase velocity.

This equation is the desired partial differential equation which is evidently satisfied by each of the *particular solutions* given in equations (33), (34), (37), (38), (39), and (40).

We shall proceed, in the following section, to deduce the same differential equations from more fundamental considerations, and shall then indicate the method by which a general solution may be deduced.

2.4 Differential Equation for the Vibration of a String.[7] It is a familiar fact that a string stretched between two fixed points, when made to vibrate, exhibits nodes and loops. The nature of the wave pattern is determined by the length of the string L, in accordance with the relation

$$\frac{n\lambda}{2} = L,$$

where n is an integral value and $\lambda/2$, which is one-half the wave length, is the distance between two consecutive nodes.

Similarly, a stretched membrane, organ pipe, or any other vibrating system exhibits definite wave patterns when vibrating, and in acoustics use is made of this fact to produce notes of definite frequencies or wave lengths. This observation that the vibrations of such systems are characterized, not by a continuously varying range of frequencies, but by a series of discrete values of these frequencies, is analogous to the spectroscopic observation that an atomic system may exist only in a series of states corresponding to a discontinuous series of values for its energy. It is this analogy, which is of a purely mathematical nature, that Schroedinger utilized in developing his "wave equation" for interpreting the behavior of electrons and atomic systems, and, therefore, before proceeding with the consideration of this wave equation, it

[7] See K. K. Darrow, *Bell System Tech. J.*, **6**, 653 (1927), for a discussion of the differential equations for vibrating systems in two and three dimensions.

is necessary to understand some aspects of the mathematical treatment of the problems of vibrating systems.

The simplest of these problems is that of a stretched string, as for instance, the wire of a piano, or the string of a violin. It is a problem involving only one coördinate in space, which is the distance along the string, and also time as a second variable, since the amplitude is a function of the time. Let us consider a stretched string, infinitely long, extended along the x-axis. Let T denote the tension along the string (that is, the force which maintains the string in a stretched condition), and ρ, the mass per unit length. To derive the differential equation for the motion of the string, we consider the forces acting on an element of length Δx, which is so short that it may be regarded as practically straight. When the string is drawn sideways (in the direction of the y-axis), these forces have components along the two axes of coördinates. Let θ denote the angle between the element and the original position of the string (along the x-axis), as shown in Fig. 7.

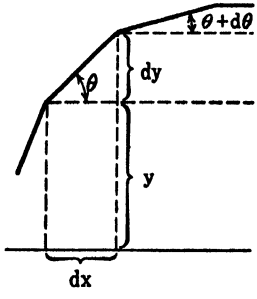


FIG. 7. Illustrating derivation of partial differential equation for vibration of a string.

The mass of the element Δx is $\rho \Delta x$, and its acceleration in the direction of the y-axis is $\partial^2 y/\partial t^2$. Hence, the force acting on the element is

$$F = \rho \frac{\partial^2 y}{\partial t^2} \Delta x. \tag{41}$$

The force F is balanced by the force arising from the difference between the tensions at the two ends of the element Δx. That is,

$$\begin{aligned} F &= T[\sin(\theta + \Delta\theta) - \sin\theta] \\ &= T[\tan(\theta + \Delta\theta) - \tan\theta], \end{aligned}$$

since, for small values of θ, $\sin\theta$ and $\tan\theta$ are approximately equal.

But
$$\tan\theta = \frac{\partial y}{\partial x},$$

and

$$\begin{aligned} \tan(\theta + \Delta\theta) - \tan\theta &= \Delta \tan\theta \\ &= \frac{d}{dx}(\tan\theta)\Delta x = \frac{\partial^2 y}{\partial x^2} \Delta x. \end{aligned}$$

Hence,
$$F = T\frac{\partial^2 y}{\partial x^2}\Delta x,$$

and, comparing this equation with (41), it follows that

$$T\frac{\partial^2 y}{\partial x^2} = \rho\frac{\partial^2 y}{\partial t^2},$$

that is,
$$\frac{\partial^2 y}{\partial x^2} = \frac{\rho}{T}\frac{\partial^2 y}{dt^2}. \tag{42a}$$

This relation is usually written in the form

$$y'' = \frac{\rho}{T}\ddot{y}, \tag{42b}$$

where primes denote differentiation with respect to x, and dots, differentiation with respect to t.

Equation (42a), or (42b), is the partial differential equation which represents a wave motion along a stretched string. It is of interest to consider the significance of the coefficient ρ/T. The constant ρ has the dimensions of mass divided by length, that is μ/l, while T has the dimensions of a force, that is $\mu l/t^2$. Thus, ρ/T has the dimensions t^2/l^2, which are identical with those of $1/u^2$ where u denotes a velocity. Therefore, we may replace ρ/T by a constant $1/u^2$ which has the additional advantage of indicating that this coefficient corresponds to a *positive* magnitude. (This is the customary method in mathematical equations of indicating such a condition and will be used quite frequently in the following sections.)

To solve the partial differential equation of the form

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{u^2}\frac{\partial^2 y}{\partial t^2}, \tag{43}$$

the classical method of procedure is that known as *solution by separation of the variables*, which is applicable provided u is a constant.

Let us assume

$$y = f(x)g(t),$$

where $f(x)$ is a function of x only, and $g(t)$ a function of t only.

Hence,

$$y'' = f''(x)g(t) = g(t)\frac{d^2 f}{dx^2},$$

and

$$\ddot{y} = f(x)\ddot{g}(t) = f(x)\frac{d^2 g}{dt^2}.$$

Since $f(x)$ and $g(t)$ are each functions of only the corresponding variable, the partial differential coefficients may be replaced by ordinary differential coefficients.

Substituting in equation (43) and dividing by $f(x)g(t)$, we obtain the relation

$$\frac{f''(x)}{f(x)} = \frac{1}{u^2}\frac{\ddot{g}(t)}{g(t)}.$$

Since the left-hand side is independent of t, while the right-hand side does not involve x, we may equate each term to an arbitrary constant, which we shall designate by $-m^2$ (to indicate that it corresponds to a negative magnitude).

We thus obtain *two ordinary differential equations*,

$$\frac{d^2f(x)}{dx^2} + m^2f(x) = 0, \tag{44}$$

and

$$\frac{d^2g(t)}{dt^2} + m^2u^2g(t) = 0. \tag{45}$$

These equations are similar to equation (6), and therefore the general solutions are of the form given in equation (7) or (12). Expressed in the exponential form, these solutions are as follows:

$$f(x) = A\epsilon^{imx} + B\epsilon^{-imx} \tag{46}$$

$$g(t) = C\epsilon^{imut} + D\epsilon^{-imut}, \tag{47}$$

where A, B, C, and D are four arbitrary constants, the values of which are determined by the "boundary" and "initial" conditions.

In the case of the string fastened at both ends, $y = 0$ for both $x = 0$ and $x = L$, where L = length of string.

Therefore, for $x = 0$,

$$f(x) = 0 = A + B.$$

Hence,

$$A = -B.$$

For $x = L$

$$f(x) = 0 = A(\epsilon^{imL} - \epsilon^{-imL})$$

$$= 2Ai \sin mL.$$

Since A is not equal to zero, $\sin mL = 0$, which means that

$$mL = n\pi; \quad m = \frac{n\pi}{L}, \tag{48}$$

where $n = 1, 2, 3$, etc.

Thus, the particular solution corresponding to equation (46) is

$$f(x) = 2Ai \sin \frac{n\pi x}{L}.$$

For $x = L/2n$, $f(x) = 2Ai$. Since $f(x)$ must be real in the case of a vibrating string, A must be a pure imaginary, and therefore $Ai = |A| = A_0$, a real quantity. Hence, the last equation should be written

$$f(x) = 2A_0 \sin \frac{n\pi x}{L}. \tag{49}$$

Turning now to the solution in equation (47), we have, as initial condition, that is for $t = 0$,

$$g(t) = 0 = C + D; \quad \text{and} \quad C = -D.$$

As for $f(x)$, we thus derive the solution

$$g(t) = 2Ci \sin mut,$$

and inserting the value of m derived in (48), this becomes (since Ci must be real)

$$g(t) = 2C_0 \sin \frac{n\pi ut}{L}.$$

Now if λ denotes the wave length, and ν the frequency, it is evident that $f(x)$ must become zero for $x = 0$, $x = \lambda/2$, $x = \lambda$, and so forth. Therefore, in equation (49) we can write

$$L = \frac{n\lambda_n}{2}, \quad \text{that is} \quad \lambda_n = \frac{2L}{n},$$

which states that L must be an integral multiple of one-half the wave length of the note emitted by the vibrating string.

Consequently,

$$f(x) = 2A_0 \sin \frac{2\pi x}{\lambda_n}$$

$$= 0 \text{ for } x = L = \frac{n\lambda_n}{2}.$$

Also it follows, since $u = \nu_n \lambda_n$, that

$$g(t) = 2C_0 \sin 2\pi \nu_n t.$$

In these last two equations, ν_n designates the frequency of the harmonic or *nth normal mode* of vibration and λ_n designates the corresponding wave length.

Thus, the particular solution of equation (43) is of the form

$$y = f(x)g(t) = 4A_0C_0 \sin\frac{2\pi x}{\lambda_n} \cdot \sin 2\pi\nu_n t, \tag{50}$$

where

$$\lambda_n = \frac{2L}{n} = \frac{\lambda}{n},$$

$$\nu_n = \frac{u}{\lambda_n} = \frac{nu}{2L} = n\nu,$$

and λ and ν are the fundamental wave length and frequency, respectively.

It will be observed that the differential equations (44) and (45) have physically significant solutions only for those discrete values of m which are defined by equation (48). These are known as *characteristic values* or *eigenvalues*, and the corresponding values of the functions $f(x)$ and $g(t)$, as defined in equation (50), are known as *characteristic functions* or *eigenfunctions*.

Corresponding to each frequency ν_n, there will be a definite vibration of an amplitude defined by the magnitude of the coefficient AC, and there will exist n loops (regions of maximum amplitude) and $(n - 1)$ nodes (regions of zero amplitude) between the ends of the wire.

It is evident that equations (46) and (47) represent a general solution of equation (43), of which equations (33), (34), etc., in section (*2*.3) are particular forms.

2.5 Schroedinger Equation for One Coördinate. The actual stimulus to the wave mechanics, as it is designated, developed by Schroedinger was derived from certain theoretical speculations of Louis de Broglie regarding the analogy between the laws of geometrical optics and those of classical dynamics.

As is well known, the laws of geometrical optics are more and more valid, the shorter the wave length of light. For light waves comparable in length with those of the object upon which they impinge, that is, for rays having a radius of curvature comparable with the wave length, we must interpret the observations from the point of view of the undulatory theory. This suggested to de Broglie the possibility that Newtonian dynamics is also an approximation which is valid for macroscopic systems, but not for atomic systems because the radius of curvature of the electronic orbit is of the same magnitude as the wave length associated with the corpuscular motion.

Let us consider, for instance, the Bohr orbit of the electron in the normal state of the hydrogen atom. The quantum condition, intro-

duced by Bohr for determining the discrete energy states, is given by the relation

$$\mu v r = \frac{h}{2\pi},$$

where v denotes the velocity of the electron in the circular orbit of radius r. If a wave length λ is associated with this motion, the circumference of the orbit must be an integral multiple of λ, and for the simplest case this multiple will be unity.

Hence,

$$2\pi r = \lambda.$$

Comparing these two equations, if follows that

$$\lambda = \frac{h}{\mu v},$$

which is the famous de Broglie relation.

As has been mentioned in Chapter I, this suggested association of wave motion with corpuscles in motion has been confirmed by the investigations on diffraction of electrons and protons. We must now consider the manner in which this observed dualism in the behavior of corpuscles has been utilized by Schroedinger in deducing his famous equation.

Let E denote the total energy of a particle, T the kinetic energy of the particle, and U the potential energy at any point in space. Then, in accordance with the law of conservation of energy

$$E = T + U$$

and

$$T = E - U.$$

For a single particle moving in a field of force, such as an electron in the hydrogen atom,

$$T = \tfrac{1}{2}\mu v^2 = E - U,$$

and therefore

$$\mu v = \sqrt{2\mu(E - U)},$$

while, since $\lambda = h/(\mu v)$,

$$\lambda = \frac{h}{\sqrt{2\mu(E - U)}}. \tag{51}$$

Since in general U is a *function of the coördinates of all the electrons* with respect to the nucleus, λ must vary from point to point in the field of force. Thus we may regard the motion of the electrons as governed by that of the associated de Broglie waves. Each electron in an atomic system follows the direction along which these waves are "refracted" in

the field of force which is due to the simultaneous presence of all the particles.

The Schroedinger (S.) equation is a partial differential equation, of the same nature as the differential equation for the vibration of a string, which indicates the manner in which the wave pattern, produced by the motion of an electron in a given field of force, varies with the space coördinates. Therefore, we introduce a function ψ, analogous to the amplitude in a physically vibrating system, which defines the "amplitude" of the wave motion and is known as the S. function or space function. Its actual physical interpretation will be considered in a subsequent section. The simplest type of problem is that in which we are dealing with the motion of a single particle of mass μ in a field of force for which the potential energy U is a function of only one coördinate x.

Hence, ψ is a function of x and t only, and we denote this by writing $\psi = \psi(x, t)$. Referring now to equation (43), which is valid for any vibrating system in which the motion occurs along only one axis of coördinate, we write the equation in the form

$$\frac{\partial^2\psi}{\partial x^2} = \frac{1}{\nu^2\lambda^2}\frac{\partial^2\psi}{\partial t^2}, \tag{52}$$

where $\nu\lambda$ takes the place of u, the velocity.

As in the solution of equation (43), we assume that

$$\psi(x, t) = \phi(x)\epsilon^{-2\pi i\nu t}, \tag{53}$$

where the function corresponding to $g(t)$ in equation (47) is assumed to be of the exponential form with a negative sign in the exponent. Evidently this satisfies the partial differential equation (52). As a matter of fact, since we are interested for the present *in the stationary wave patterns*, we may postpone for a subsequent chapter any further consideration of the exact function by which to represent the variation in $\psi(x, t)$ with t.

Incorporating the relation for ψ, assumed in equation (53), into the differential equation (52), we derive the relations

$$\frac{\partial\psi}{\partial t} = -2\pi i\nu\phi \cdot \epsilon^{-2\pi i\nu t}$$

$$\frac{\partial^2\psi}{\partial t^2} = -4\pi^2\nu^2\phi \cdot \epsilon^{-2\pi i\nu t},$$

and

$$\frac{\partial^2\psi}{\partial x^2} = \frac{d^2\phi}{dx^2} \cdot \epsilon^{-2\pi i\nu t}.$$

In these equations ϕ is used instead of $\phi(x)$.

Substituting the last two relations into (52), the result is the ordinary differential equation

$$\frac{d^2\phi}{dx^2} + \frac{4\pi^2\phi}{\lambda^2} = 0.$$

In consequence of equation (51) this equation assumes the form

$$\frac{d^2\phi}{dx^2} + \frac{8\pi^2\mu(E - U)\phi}{h^2} = 0. \tag{54}$$

This is the form of the S. equation for one coördinate, and it will be observed that, since U is a function of x, the equation cannot be solved as readily as the differential equations considered in the previous sections. The mathematical technic of solution thus depends upon the nature of the function $U = U(x)$.

2.6 Motion of Electrons in Absence of Field of Force. In absence of a field, that is, for $U = 0$, the S. equation assumes the form

$$\frac{d^2\phi}{dx^2} + \frac{8\pi^2\mu E\phi}{h^2} = 0. \tag{55}$$

Since the coefficient of ϕ is a constant, this equation has the same form as equation (44). Thus, the solution is

$$\phi = \phi(x) = Ae^{i\alpha x} + Be^{-i\alpha x}, \tag{56}$$

where $\alpha^2 = 8\pi^2\mu E/h^2$, and A and B are arbitrary constants.

As emphasized in a previous section, this equation has physical significance only if $\alpha = 2\pi/\lambda$, that is, if

$$2\pi\frac{\sqrt{2\mu E}}{h} = \frac{2\pi}{\lambda},$$

or

$$\lambda = \frac{h}{\sqrt{2\mu E}}, \tag{57}$$

which is de Broglie's relation.

What is the physical interpretation of the function ϕ? In equation (37) it was shown that the expression

$$e^{i(\alpha x - \omega t)} = e^{i\alpha x}\, e^{-2\pi i\nu t}$$

represents a simple harmonic wave for which the direction of propagation is from left to right, while the expression

$$e^{-i(\alpha x + \omega t)} = e^{-i\alpha x}\, e^{-2\pi i\nu t}$$

represents a simple harmonic wave motion from right to left.

Thus, if in equation (56), we wish to designate a wave motion propagated from left to right, we can indicate this by putting $B = 0$. To obtain an interpretation of ϕ we proceed as follows:

Multiplying ϕ by its complex conjugate, the result is

$$\phi\bar{\phi} = A\bar{A}\epsilon^{i\alpha x} \cdot \epsilon^{-2\pi i\nu t} \cdot \epsilon^{-i\alpha x} \cdot \epsilon^{2\pi i\nu t}$$
$$= A\bar{A} = |A|^2. \tag{58}$$

That is $\phi\bar{\phi}$ is a *real* magnitude. In quantum mechanics $\phi\bar{\phi}dx$ is regarded as the probability of occurrence of an electron in the element of distance dx at the point x. Since $\phi\bar{\phi}$ is a constant in the present instance, we may interpret this magnitude as designating the *number of electrons per unit length.* The fact that this linear density is constant means that the electrons are *uniformly distributed* at all points along the x-coördinate from $-\infty$ to $+\infty$. This result is in agreement with the deduction from the Principle of Indeterminacy. For an electron having a definite momentum $p = \mu v = \sqrt{2\mu E}$, the associated wave motion is represented by a monochromatic (or unifrequentic) wave extending from $-\infty$ to $+\infty$, and the position of the electron is completely undetermined. There is an equal probability for the occurrence of the electron anywhere along the infinite wave train.

The case $A = \pm B$ evidently corresponds to a *stationary* de Broglie wave. For

$$\phi = A(\epsilon^{i\alpha x} \pm \epsilon^{-i\alpha x})\epsilon^{-2\pi i\nu t}$$
$$= 2A \cdot \cos \alpha x \cdot \epsilon^{-2\pi i\nu t}$$

or

$$= 2Ai \cdot \sin \alpha x \cdot \epsilon^{-2\pi i\nu t},$$

depending upon the particular physical conditions to be satisfied. In this case

$$\left.\begin{aligned} \phi\bar{\phi} &= 4A^2 \cos^2 \alpha x \\ \text{or} \qquad &= 4A^2 \sin^2 \alpha x \end{aligned}\right\}, \tag{59}$$

and therefore the value of the distribution functions along the x-axis varies periodically between the values 0 and $4A^2$. Again, we interpret the expression $\phi\bar{\phi}dx$ as representing the relative probability of occurrence of electrons in the element dx at the point x.

From these considerations it is possible to perceive the reasons for choosing the exponential functions $\epsilon^{i\alpha x}$ and $\epsilon^{-i\alpha x}$ for representing the motion of a unidirectional stream of electrons which have a *definite velocity.* It is only by means of these functions that it is possible to derive a value of $\phi\bar{\phi}$ which is *constant for all points along the direction of*

motion, and which is in accordance with the observation that in such an electron stream the instantaneous location of any one electron is completely indeterminable.

If we attempt to use the sine or cosine functions, the value of $\phi\bar{\phi}$ thus derived turns out to be a function of x, as shown in (59), which would indicate the possibility of observing that the electron density varies at different points along the direction of motion in accordance with the variation of $\cos \alpha x$ or $\sin \alpha x$. Hence we are justified in using the latter representations only where observation indicates the existence of a corresponding density distribution along the axis of x.

We shall now consider the interpretation of ν and u for a de Broglie wave motion. Since it is not possible to measure ν for such a wave by any experimental method (the only magnitude which may be determined experimentally is the wave length), we assume that the frequency is defined by the quantum relation $\nu = E/h$, where E is the total energy. According to de Broglie, the value of E that should be used in calculating ν is that derived on the basis of the special theory of relativity, that is, $E + \mu c^2$, where c = velocity of light. However, in any observations, it is only energy differences that are actually measured, and, therefore, it is immaterial which value of the total energy is used. Owing to this indefiniteness in the absolute value of E, it follows that the value of $u = \nu\lambda$ also cannot be determined by any experimental method.

In treatises on quantum mechanics, u is designated as phase velocity and a distinction is drawn between this quantity, which cannot be observed experimentally, and the so-called group velocity, which is identical with the experimentally observed corpuscular velocity v. However, to the writer it seems that altogether too much emphasis has been laid on this topic, at least in an introduction to quantum mechanics. As a matter of fact, it is meaningless to speak of a velocity u, and a frequency ν, as if they were observable magnitudes. For this would imply a physical phenomenon in which something is, as it were, in a state of vibration. But actually we find that in many problems ϕ is a function of more than three coördinates. Thus, in the problem of the helium atom, ϕ for the system is a function of the coördinates of each electron and, therefore, represents a vibration in a space of six dimensions. Obviously, "phase velocity" can have no physical significance in such a case, although it is possible to calculate a value of $\phi\bar{\phi}dxdydz$, which gives the relative probability of locating one of the electrons in the element of volume $dxdydz$ at any point in the space surrounding the nucleus.

As N. F. Mott has stated,[8] in referring to the function ϕ, "The wave

[8] N. F. Mott, "An Outline of Wave Mechanics," p. 51.

function is, so to speak, just a convenient shorthand. The waves are not waves in any medium. There are no waves accompanying the electron, until we have observed the electron. Then we make use of the wave representation to embody the results of an observation. The wave equation tells us what may be deduced, from our observations, about the future position and velocity of an electron."

2.7 Operator Method of Deriving the Schroedinger Equation. This divorcing of the S. equation from any physical interpretation is demonstrated by considering an alternative method of deriving the equation which involves no mention whatever of a de Broglie wave length.

Let us refer once more to the equation for the total energy of a system, as derived from classical mechanics. We write this, for a single particle, in the form

$$E = \tfrac{1}{2}\mu v^2 + U(x)$$

or

$$\frac{p^2}{2\mu} + U(x) - E = 0, \tag{60}$$

where $p = \mu v$ = momentum.

This equation gives the relation for the energy in the so-called Hamiltonian form, that is, as a function of coördinates and corresponding momenta. This is indicated symbolically by writing $E = H(p, q)$, where q is used instead of x to designate the coördinate.[9]

We now replace p by a differential *operator*, which we assume to be of the form

$$p = \frac{h}{2\pi i} \cdot \frac{\partial}{\partial q}, \tag{61}$$

where the partial differential coefficient refers to the more general case in which E is a function of several coördinates, and of their corresponding momenta. In the present case, the partial differential is synonymous with the ordinary differential, so that the symbol ∂ may be replaced by d. Thus, we obtain the operator

$$-\frac{h^2}{4\pi^2} \cdot \frac{1}{2\mu} \cdot \frac{d^2}{dx^2} + U(x) - E.$$

Applying this operator to the function ϕ, the result is

$$\left\{ \frac{d^2}{dx^2} + \frac{8\pi^2\mu}{h^2} (E - U) \right\} \phi = 0, \quad \cdot$$

[9] In Chapter IV this method of expressing the total energy is discussed more fully.

that is,

$$\frac{d^2\phi}{dx^2} + \frac{8\pi^2\mu}{h^2}(E - U)\phi = 0, \tag{62}$$

which is the S. equation.

This method of deriving the S. equation involves a step which is very fundamental in the new quantum mechanics. Instead of speaking of observable magnitudes which characterize a given state of a system, we consider rather the kind of operations which may be performed upon the function, which is characteristic of that state. To each observable in the classical mechanics formulation there corresponds a certain operator, and the results of observations on the system are predicted on the basis of deductions derived by operating on the characteristic function of a given state.

This use of the so-called algebra of operators is characteristic of the new quantum mechanics, but it is worth while pointing out that, even in the more elementary types of mathematics, the use of operators is quite general. Thus, when we write y^3 we indicate the successive operations $y \cdot y \cdot y$, and the notation $\sqrt{y}$ indicates the operation of taking the square root. Other operators are *log* and *sin* (*cos*, *cosh*, *sinh*, etc.). The notation $\sin\theta$ indicates a certain operation to be performed on the angle θ. Similarly, $d/d\theta = D_\theta$ corresponds to another kind of operation. Thus, the expression $D_\theta \sin\theta$ indicates the following sequence of operations: (1) finding the sine of θ, and (2) determining the rate of change in $\sin\theta$ with change in θ. It will be observed from this example that the *order* in which the operators are given is of fundamental importance.

Thus, although operators may be treated as algebraic symbols, that is, we may multiply and also (usually) divide by them, as if they were actual magnitudes, they are not in general *commutative*. That is, in general, if α and β are two different operators

$$\alpha\beta \neq \beta\alpha$$

and $\alpha\beta - \beta\alpha$ is known as the *commutator* of the operators α and β. As an illustration let us consider the commutator

$$Dx - xD$$

of the operators D and x. The result of operating on unity is

$$(Dx - xD)1 = \frac{d}{dx}x \cdot 1 - x\frac{d}{dx}1$$

$$= 1 - 0 = 1.$$

Again, if we operate on x, the result is

$$(Dx - xD)x = \frac{d}{dx}(x^2) - x\frac{d}{dx}x$$

$$= 2x - x = x,$$

and similarly

$$(Dx - xD)x^n = (n+1)x^n - nx^n$$

$$= x^n.$$

Thus, the commutator of the operators D and x is unity.

In the case of the operator p as defined in equation (61), it follows that

$$pq - qp = \frac{h}{2\pi i}\frac{\partial}{\partial q}q - q\frac{h}{2\pi i}\frac{\partial}{\partial q}1$$

$$= \frac{h}{2\pi i}. \tag{63}$$

That is, p and q are *non-commutative*, and the commutator of these operators, in the order p, q, is $h/(2\pi i)$.

This result represents the analog, in the new mechanics, of the Bohr-Sommerfeld quantizing condition for the determination of electronic orbits, which was stated in the form

$$\oint p_k q_k dq_k = nh,$$

where n is an integral value, p_k and q_k are "canonically conjugated" variables (see Chapter IV), and the circle on the integral sign indicates that the integration is to be carried out over the whole orbit (periodic path).

From equation (63) it is possible to deduce the Principle of Indeterminacy, and it is also clear, since p_k and q_k do not commute, that these variables cannot have the same relative significance in the new theory as they had in classical mechanics.

It is of interest to point out further in what manner use may be made of the operator concept in deducing a physical result.

In section 3 we considered the interpretation of the solutions

$$\phi_1 = \epsilon^{i\alpha x}$$

$$\phi_2 = \epsilon^{-i\alpha x}$$

for the motion of a homogeneous beam of electrons in absence of a field.

Applying the operator p to each side of the first of these relations, we obtain the result

$$p\phi_1 = \frac{h}{2\pi i} \cdot \frac{d}{dx} (\epsilon^{i\alpha x})$$

$$= \frac{hi\alpha}{2\pi i} \cdot \epsilon^{i\alpha x} = \sqrt{2\mu E} \cdot \phi_1, \tag{64}$$

since $\alpha = (2\pi/h)\sqrt{2\mu E}$.

That is, the result of operating on the function ϕ_1 with the operator p leads to a relation in which the coefficient of ϕ_1 is a constant. This means, physically, that if an experiment is performed on the beam of electrons to determine their momentum the observation will lead to a definite value given by the relation

$$p = \sqrt{2\mu E}. \tag{65}$$

The positive sign before the radical indicates that p and therefore $v = dx/dt$ is positive. That is, the electrons are propagated from left to right.

Similarly it may be shown that in case of ϕ_2,

$$p = -\sqrt{2\mu E},$$

that is, the wave is propagated from right to left.

The case

$$\phi = \epsilon^{i\alpha x} \pm \epsilon^{-i\alpha x}$$

$$= 2 \cos \alpha x$$

or

$$= 2i \sin \alpha x$$

leads to the relation

$$p\phi = -\frac{h\alpha}{\pi i} \cdot \sin \alpha x,$$

in the first case, that is,

$$p(\cos \alpha x) = -\frac{h\alpha}{2\pi i} \cdot \sin \alpha x, \tag{66}$$

and in the second case, to the relation

$$p(\sin \alpha x) = \frac{h\alpha}{2\pi} \cdot \cos \alpha x. \tag{67}$$

It will be noted that the last two equations differ from (64) where the result of operating on the function with p led to a product of a constant and the function which it was operated upon. In the case of equation (64) we were led to the conclusion, indicated by equation (65), that the momentum has a precise value. On the other hand, equations (66) and (67) indicate that we may no longer draw a conclusion of this nature when dealing with the function for a stationary wave pattern. In fact, any attempt to determine the direction of motion of electrons in the latter case will not lead to repeatable results.

COLLATERAL READING

References on Mathematics

1. MELLOR, J. W., "Higher Mathematics for Students of Chemistry and Physics, Longmans, Green & Co., London, 1929. Contains a splendid section o differential equations, hyperbolic functions, etc.
2. PHILLIPS, H. B., "Differential Equations," John Wiley & Sons, New York 1924. A very elementary treatment.
3. HITCHCOCK, F. L., and ROBINSON, C. S., "Differential Equations in Applie Chemistry," John Wiley & Sons, New York, 1936. Good chapter on Fourier's equation.
4. HUMPHREY, D., "Advanced Mathematics for Students of Physics and Engineering," Oxford University Press, 1929. Two volumes which contain a great deal about differential equations, Fourier series, and spherical harmonics.
5. HOUSTOUN, R. A., "An Introduction to Mathematical Physics," Longmans, Green & Co., London, 1929. An extremely useful work of reference on problems of potential, wave motion, Fourier series, and electromagnetic theory.
6. DOHERTY, R. E., and KELLER, E. G., "Mathematics of Modern Engineering," Vol. 1, John Wiley & Sons, New York, 1936. Comprehensive discussion of the solution of differential equations, Fourier series, dimensional analysis, and vector analysis.
7. HOUSTON, W. V., "Principles of Mathematical Physics," McGraw-Hill Book Co., New York, 1934. More advanced than R. A. Houstoun's book and also discusses topics of interest in connection with classical and quantum mechanics.
8. MADELUNG, E., "Die mathematischen Hilfsmittel des Physikers," Julius Springer, Berlin, 1925. An easier treatise to read than the two preceding works.
9. COURANT, R., and HILBERT, D., "Methoden der mathematischen Physik, I," Julius Springer, Berlin, 1924. The most comprehensive book on the subject, but rather difficult for students who have not gone beyond differential equations.
10. FRANK, P., and MISES, R. v., "Die Differential- und Integralgleichungen der Mechanik und Physik," Friedr. Vieweg & Sohn, Akt.-Ges., Braunschweig, 1930. Two volumes.
11. Handbuch der Physik, III. Mathematische Hilfsmittel in der Physik, Julius Springer, 1928. Contains chapters on a number of topics. This volume and the treatise by Frank and v. Mises are in the same class as Courant and Hilbert.

CHAPTER III

PROBLEMS OF POTENTIAL BARRIERS

The logic of the S. method and its applications may be understood most readily by a consideration of a very important class of problems — those of the transmission of electrons through potential barriers. The recognition that it is possible, on the basis of the new point of view, for electrons, possessing a given kinetic energy E, to penetrate into a region for which the potential energy U exceeds E — this conclusion, derived on the basis of the S. equation, has enabled us to interpret such observations as those on radioactive disintegration and the emission of electrons from cold cathodes in the presence of high field strengths.

3.1 Reflection of Electrons at Semi-Infinite Potential Barriers. The simplest problem, which at the same time illustrates the difference between the conclusions which would be drawn from arguments based on classical mechanics and those derived on the basis of the S. equation, is that of a beam of electrons incident on a boundary, at which the value of the potential energy increases abruptly from 0 to U.

Let us consider a homogeneous beam of electrons of kinetic energy E moving along the x-axis from left to right, and incident at $x = 0$ on a boundary at which there is a retarding potential $U = U_0$ for all values of $x \geqslant 0$. Two distinct cases arise here: one, illustrated in Fig. 8, for which $E > U_0$; the other, illustrated in Fig. 9, for which $E < U_0$.

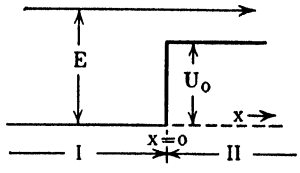


FIG. 8. Semi-infinite potential barrier; Case I. $E > U_0$.

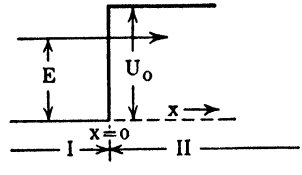


FIG. 9. Semi-infinite potential barrier; Case II. $E < U_0$.

In Case I ($E > U_0$), according to classical mechanics, a particle passing from left to right would merely be slowed up in crossing the boundary at $x = 0$. On the other hand, in Case II ($E < U_0$), there would be complete reflection. While quantum mechanics also shows that there is a difference in the behavior of the particles in these two

cases, certain conclusions are deduced which are different from those predicted by classical mechanics.

Case I. $E > U_0$.

Let us designate the region to the left of $x = 0$, $(U = 0)$ by I, that to the right of $x = 0$, $(U = U_0)$ by II. Let ϕ_1 and ϕ_2 denote the wave functions for each of these regions. Each function will satisfy a S. equation, which will be of the form

$$\frac{d^2\phi_1}{dx^2} + \frac{8\pi^2\mu}{h^2} E\phi_1 = 0 \tag{1}$$

for region I, and

$$\frac{d^2\phi_2}{dx^2} + \frac{8\pi^2\mu}{h^2} (E - U_0)\phi_2 = 0 \tag{2}$$

for region II.

The solutions of these equations are, respectively,

$$\phi_1 = A\epsilon^{i\alpha x} + B\epsilon^{-i\alpha x} \tag{3}$$

$$\phi_2 = C\epsilon^{i\beta x} + D\epsilon^{-i\beta x}, \tag{4}$$

where

$$\alpha^2 = \frac{8\pi^2\mu E}{h^2}; \quad \text{and} \quad \beta^2 = \frac{8\pi^2\mu}{h^2} (E - U_0).$$

Evidently, $\alpha = 2\pi/\lambda_1$, and $\beta = 2\pi/\lambda_2$, where λ_1 and λ_2 are the de Broglie wave lengths in each region.

Now we introduce the boundary conditions which have to be satisfied by ϕ_1 and ϕ_2, in order that the solutions shall be physically significant. The first condition is that at $x = 0$

$$\phi_1 = \phi_2. \tag{5}$$

That is, there can be no discontinuity in the *value* of the wave function at the boundary.

The second condition is that there shall be no discontinuity in the *slope*, that is, at $x = 0$,

$$\frac{d\phi_1}{dx} = \frac{d\phi_2}{dx}. \tag{6}$$

These two conditions are necessary in order to insure the continuity in the nature of the wave function as it crosses the boundary at $x = 0$. These are the conditions which are used in classical theory in dealing with the behavior of waves at a boundary. Furthermore, we know

from the physics of the problem that the incident beam of electrons may give rise to two streams, one of which represents the electrons that are transmitted from region I into region II, and the other represents the electrons reflected at $x = 0$. From the considerations advanced in the previous chapter, it follows that the incident beam must be represented by $A\epsilon^{i\alpha x}$, the reflected beam by $B\epsilon^{-i\alpha x}$, and the transmitted beam by $C\epsilon^{i\beta x}$. Since there are no electrons moving from right to left in region II, $D = 0$.

Introducing these conclusions into (3) and (4), along with the two boundary conditions, equations (5) and (6), we obtain the relations:

$$A + B = C$$

$$\alpha(A - B) = \beta C.$$

Hence,

$$B = \frac{\alpha - \beta}{\alpha + \beta} \cdot A$$

$$C = \frac{2\alpha}{\alpha + \beta} \cdot A,$$

and therefore,

$$\phi_1 = A\left(\epsilon^{i\alpha x} + \frac{\alpha - \beta}{\alpha + \beta} \cdot \epsilon^{-i\alpha x}\right) \tag{7}$$

$$\phi_2 = A \cdot \frac{2\alpha}{\alpha + \beta} \cdot \epsilon^{i\beta x}. \tag{8}$$

Thus, if the concentration, that is the number of electrons per unit distance, is equal, in the case of the incident beam, to A^2, then it follows that

(a) the concentration in reflected beam $= A^2\left(\frac{\alpha - \beta}{\alpha + \beta}\right)^2$

(b) the concentration in transmitted beam $= A^2 \cdot \frac{4\alpha^2}{(\alpha + \beta)^2}$.

To obtain the coefficient of transmission T and that of reflection R, we must compare the "currents" in the three beams. The *rate* at which the electrons strike the potential barrier is equal to the sum of the rates of transmission and reflection.

Now current = (concentration) × (velocity).
For region I, the velocity v_1 is given by

$$v_1 = \sqrt{\frac{2E}{\mu}} = \frac{h\alpha}{2\pi\mu},$$

while for region II,

$$v_2 = \sqrt{\frac{2(E - U_0)}{\mu}} = \frac{h\beta}{2\pi\mu}.$$

Therefore,

$$R = \left(\frac{\alpha - \beta}{\alpha + \beta}\right)^2 \frac{\alpha}{\alpha} = \frac{\alpha^2 - 2\alpha\beta + \beta^2}{(\alpha + \beta)^2}, \tag{9}$$

$$T = \frac{4\alpha^2}{(\alpha + \beta)^2} \frac{\beta}{\alpha} = \frac{4\alpha\beta}{(\alpha + \beta)^2}, \tag{10}$$

and evidently $R + T = 1$, which satisfies the physical requisites. Replacing α and β by the corresponding values in terms of E and U_0, it is readily shown that

$$R = \left(\frac{1 - \sqrt{\frac{E - U_0}{E}}}{1 + \sqrt{\frac{E - U_0}{E}}}\right)^2. \tag{11}$$

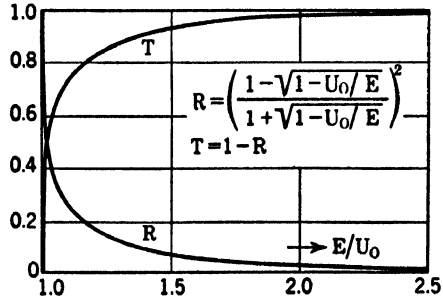


FIG. 10. Coefficients of reflection R and transmission T for electrons as a function of E/U_0.

Table I gives values of R for different values of U_0/E as derived by means of equation (11), and Fig. 10 shows plots of T and R as functions of E/U_0.

TABLE 1

$\frac{U_0}{E}$	R	$\frac{E}{U_0}$
0.1	0.0007	10.0
0.2	0.0034	5.0
0.4	0.0161	2.5
0.5	0.0296	2.0
0.6	0.0506	1.67
0.7	0.0852	1.43
0.8	0.1459	1.25
0.9	0.270	1.11
1 0	1.000	1 0

Classical mechanics predicts the result $R = 0$ for all values of $E \geqslant U_0$; quantum mechanics states that even for $E = 2.5\ U_0$, 1.61 per cent of the incident electrons are reflected.

Case II. $E < U_0$.

As in the preceding case we consider the two S. equations:

$$\frac{d^2\phi_1}{dx^2} + \frac{8\pi^2\mu}{h^2} \cdot E\phi_1 = 0, \tag{12}$$

and

$$\frac{d^2\phi_2}{dx^2} + \frac{8\pi^2\mu}{h^2}(E - U_0)\phi_2 = 0. \tag{13a}$$

The solution of (12) is, as before,

$$\phi_1 = A\epsilon^{iax} + B\epsilon^{-iax}. \tag{14}$$

In equation (13*a*) U_0 is greater than E and, hence, the momentum and corresponding de Broglie wave length are imaginary. If we write the equation in the form

$$\frac{d^2\phi_2}{dx^2} - \frac{8\pi^2\mu}{h^2}(U_0 - E)\phi_2 = 0, \tag{13b}$$

where $U_0 - E$ is a *positive quantity*, the general solution may be written, by analogy with equation (*2*.20), as

$$\phi_2 = C\epsilon^{\beta x} + D\epsilon^{-\beta x}, \tag{15}$$

where

$$\beta^2 = \frac{8\pi^2\mu}{h^2}(U_0 - E). \tag{16}$$

Equation (15) represents the sum of a function $C\epsilon^{\beta x}$ which increases exponentially with increase in x, and another function $D\epsilon^{-\beta x}$ which decreases exponentially in the same region. Now purely physical considerations show that if any electrons actually penetrate into region II, as is indicated by the solution ϕ_2, then the concentration of these electrons must decrease rapidly with increase in x. Hence, we must put $C = 0$.

Applying the boundary conditions stated in equations (5) and (6), we deduce the relations:

$$A + B = D,$$

$$i\alpha(A - B) = -\beta D.$$

That is,

$$A - B = \frac{i\beta}{\alpha} \cdot D.$$

Therefore,

$$A = \frac{D}{2}\left(1 + \frac{i\beta}{\alpha}\right)$$

$$B = \frac{D}{2}\left(1 - \frac{i\beta}{\alpha}\right).$$

That is, A and B are complex conjugate quantities.

Hence,

$$\phi_1 = \frac{D}{2}\left(1 + \frac{i\beta}{\alpha}\right)\epsilon^{i\alpha x} + \frac{D}{2}\left(1 - \frac{i\beta}{\alpha}\right)\epsilon^{-i\alpha x}$$

$$= \frac{D}{2}(\epsilon^{i\alpha x} + \epsilon^{-i\alpha x}) - \frac{D}{2i} \cdot \frac{\beta}{\alpha}(\epsilon^{i\alpha x} - \epsilon^{-i\alpha x})$$

$$= D \cos \alpha x - D \cdot \frac{\beta}{\alpha} \sin \alpha x$$

$$= D\frac{\sqrt{\alpha^2 + \beta^2}}{\alpha}\left(\cos \alpha x \cdot \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}} - \sin \alpha x \cdot \frac{\beta}{\sqrt{\alpha^2 + \beta^2}}\right)$$

$$= D\sqrt{1 + \frac{\beta^2}{\alpha^2}} \cos(\alpha x + \delta), \tag{17}$$

where

$$\cos \delta = \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}}; \quad \sin \delta = \frac{\beta}{\sqrt{\alpha^2 + \beta^2}}.$$

The average concentration per unit length (C_0) in the region $x < 0$ is obtained thus:

Since the de Broglie wave length is given by the relation

$$\lambda = \frac{h}{\sqrt{2\mu E}} = \frac{2\pi}{\alpha},$$

$$C_0 = \frac{2}{\lambda}\int_0^{\frac{\lambda}{2}} \phi_1^2\, dx = \frac{1}{\pi}\int_0^{\frac{\pi}{\alpha}} \phi_1^2\, \alpha dx.$$

Let $x = \theta/\alpha$. Then we can write $\cos^2 \alpha x = \cos^2 \theta$, and the limits of integration are 0 and π. Since

$$\int_0^{\pi} \cos^2 \theta d\theta = \frac{\pi}{2},$$

$$C_0 = \frac{D^2}{2}\left(1 + \frac{\beta^2}{\alpha^2}\right) = 2 \mid A \mid^2 = 2 \mid B \mid^2. \tag{18}$$

Equation (15), with $C = 0$ and D real, shows that there is a definite probability for the occurrence of the electrons in the region $U > E$. The *total probability* of the occurrence of electrons in this region is

$$P = \int_0^{\infty} \phi_2^2 dx = \int_0^{\infty} D^2 \epsilon^{-2\beta x}\, dx = \frac{D^2}{2\beta}. \tag{19}$$

The ratio P/C_0 gives, as E. U. Condon[1] terms it, "a kind of mean depth of penetration of the particles into the non-classical region." It is evidently

$$\frac{P}{C_0} = \frac{\alpha^2}{\beta(\alpha^2 + \beta^2)}, \tag{20}$$

and varies from 0 for $E = 0$ to ∞ as E tends to become equal to U_0. The particles, of course, do not stay in region II indefinitely but ultimately return to region I. Also it is evident that with the increasing value of the ratio U_0/E (corresponding to increasing values of β/α) the relative probability of penetration, as defined by equation (20), decreases until, for $\beta = \infty$, this probability becomes zero, that is, there is no penetration of particles into the "forbidden" region.

In this case, since $\phi_1 = 0$ for $x = 0$, it follows that

$$\phi_1 = A(\epsilon^{i\alpha x} - \epsilon^{-i\alpha x})$$
$$= 2Ai \sin \alpha x.$$

[1] E. U. Condon, *Rev. Modern Phys.*, **3**, 43 (1931).

That is, the eigenfunction represents two streams of electrons of equal intensity but in opposite directions. Thus, in the case $E < U_0$, the result obtained for R, the reflection coefficient, is the same as in classical mechanics. There is, however, this difference. Whereas the classical treatment states that penetration of electrons into the region $U_0 > E$ is forbidden, quantum mechanics states that there exists a definite probability for the occurrence of the phenomenon and that this probability becomes vanishingly small as the ratio β/α tends to infinity.

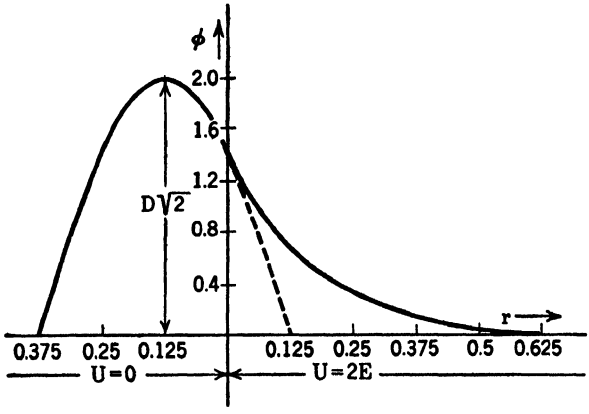


FIG. 11. Form of eigenfunction for penetration in region $E < U$.

Figure 11 shows a plot of ϕ_1 and ϕ_2 for the case $U_0 = 2E$. These curves were derived as follows:

$$\alpha x = \frac{2\pi x}{\lambda} = 2\pi r,$$

where

$$\lambda = \frac{h}{\sqrt{2\mu E}} \quad \text{and} \quad r = \frac{x}{\lambda}.$$

Also

$$\beta = \alpha \sqrt{\frac{U_0}{E} - 1}.$$

Hence, in this case,

$$\beta x = 2\pi r \sqrt{\frac{U_0}{E} - 1} = 2\pi r.$$

Thus $\cos \delta = \sin \delta = 1/\sqrt{2}$; $\delta = 2\pi/8$, and the corresponding value of r is 0.125.

It follows that

$$\phi_1 = D\sqrt{2} \cdot \cos 2\pi(r + 0.125),$$

while

$$\phi_2 = D\epsilon^{-2\pi r}.$$

For

$$r = 0; \quad \phi_1 = \phi_2 = D,$$

and

$$\frac{d\phi_1}{dr} = -2\pi D = \frac{d\phi_2}{dr}.$$

It will also be observed that $C_0 = D^2$ is the concentration per unit length in the region $U = 0$.

3.2 Electrons between Semi-Infinite Potential Barriers. A problem which may be considered at this stage is that of an electron in a potential "box."[2] This is illustrated in Fig. 12, for the case in which $U = 0$ for the region $x = -a$ to $x = +a$, and has the value $U > E$ at all points (extending to $-\infty$ and $+\infty$) outside this region. What will be the behavior of an electron confined in such a region which we assume to extend to infinity along the y- and z-coördinates?

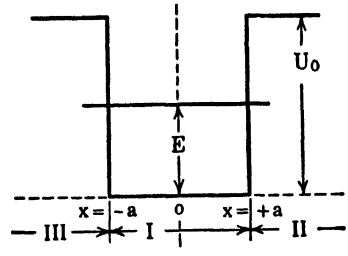


FIG. 12. Illustrating behavior of electron in potential "box."

The problem is of importance for at least two reasons. Firstly, the behavior of an electron confined between two potential barriers, such as those assumed in the present case, must present similarities to the behavior of an electron bound by the electrostatic field around a positively charged nucleus. Secondly, the problem resembles, in its simplest aspects, that of electrons in a metal. In this case, as is well known, the kinetic energy of the electrons (E) begins to exceed the work function (V_0) for emission, only at higher temperatures, and the increase of emission with temperature corresponds to an increase in the number of electrons in the metal which possess a kinetic energy greater than or equal to the work function.

In the case shown in Fig. 12 we have to consider three regions: (1) that extending from $x = -a$ to $x = +a$; (2) that from $x = +a$ to $x = \infty$; (3) that from $x = -a$ to $x = -\infty$. Let us write

$$\alpha^2 = \frac{8\pi^2\mu E}{h^2}; \quad \beta^2 = \frac{8\pi^2\mu}{h^2}(U_0 - E),$$

where, as in the previous case, since $U_0 > E$, β^2 is a positive quantity.

[2] The solution of this problem is given by J. Frenkel, "Einführung in die Wellenmechanik," pp. 52–55; "Wave Mechanics," Vol. I.

The solutions of the S. equation are given by

$$\phi_{\mathrm{I}} = A\epsilon^{i\alpha x} + B\epsilon^{-i\alpha x}$$
$$\phi_{\mathrm{II}} = C\epsilon^{-\beta x},$$

which makes ϕ_{II} vanish for large positive values of x, and

$$\phi_{\mathrm{III}} = D\epsilon^{\beta x},$$

which makes ϕ_{III} vanish for large negative values of x.

Applying the continuity relations for $x = a$, we have

$$\phi_{\mathrm{I}} = \phi_{\mathrm{II}}$$
$$\frac{d\phi_{\mathrm{I}}}{dx} = \frac{d\phi_{\mathrm{II}}}{dx}.$$

From these relations and from similar relations for $x = -a$, we obtain the four equations

$$A\epsilon^{i\alpha a} + B\epsilon^{-i\alpha a} = C\epsilon^{-\beta a} \quad \text{(i)}$$
$$A\epsilon^{-i\alpha a} + B\epsilon^{i\alpha a} = D\epsilon^{-\beta a} \quad \text{(ii)}$$
$$i\alpha(A\epsilon^{i\alpha a} - B\epsilon^{-i\alpha a}) = -\beta C\epsilon^{-\beta a} \quad \text{(iii)}$$
$$i\alpha(A\epsilon^{-i\alpha a} - B\epsilon^{i\alpha a}) = \beta D\epsilon^{-\beta a} \quad \text{(iv)}$$

which is a system of four linear homogeneous equations in the unknowns A, B, C, and D.

From (i) and (ii) by addition,

$$2(A + B)\cos \alpha a = (C + D)\epsilon^{-\beta a}. \quad \text{(v)}$$

From (iii) and (iv) by subtraction,

$$2\alpha(A + B)\sin \alpha a = \beta(C + D)\epsilon^{-\beta a}. \quad \text{(vi)}$$

Hence,

$$\tan \alpha a = \frac{\beta}{\alpha}. \quad (21)$$

Again, from (i) and (ii) by subtraction,

$$2i(A - B)\sin \alpha a = (C - D)\epsilon^{-\beta a}, \quad \text{(vii)}$$

and from (iii) and (iv) by addition,

$$2i\alpha(A - B)\cos \alpha a = -\beta(C - D)\epsilon^{-\beta a}. \quad \text{(viii)}$$

Hence,

$$\tan \alpha a = -\frac{\alpha}{\beta}. \quad (22)$$

If (21) is valid, then it follows from (vii) and (viii) that

$$A = B;\ C = D = 2A\epsilon^{\beta a}\cos \alpha a. \quad (23)$$

From (22) and (v) and (vi) it follows that

$$A = -B;\ C = -D = 2Ai\epsilon^{\beta a} \sin \alpha a. \tag{24}$$

We thus obtain *two sets of values of the constants, which satisfy the boundary conditions of the problem.* The corresponding eigenfunctions in the case of equations (21) and (23) are

$$\left.\begin{aligned} \phi_{\mathrm{I}} &= 2A \cos \alpha x && (-a \leqq x \leqq a) \\ \phi_{\mathrm{II}} &= 2A \cos \alpha a \cdot \epsilon^{-\beta(x-a)} && (x > a) \\ \phi_{\mathrm{III}} &= 2A \cos \alpha a \cdot \epsilon^{\beta(x+a)} && (x < -a) \end{aligned}\right\}. \tag{25}$$

While in the case of equations (22) and (24), the eigenfunctions are

$$\left.\begin{aligned} \phi_{\mathrm{I}} &= 2iA \sin \alpha x \\ \phi_{\mathrm{II}} &= 2iA \sin \alpha a \cdot \epsilon^{-\beta(x-a)} \\ \phi_{\mathrm{III}} &= -2iA \sin \alpha a \cdot \epsilon^{\beta(x+a)} \end{aligned}\right\}. \tag{26}$$

Equations (21) and (25) thus represent one type of vibration, in which for $x = 0, \phi_{\mathrm{I}} = 2A$, and for $x = \pm a, \phi_{\mathrm{I}} = 2A \cos \alpha a$, while (22) and (26) represent another type of vibration in which, for $x = 0$, $\phi_{\mathrm{I}} = 0$, and for $x = +a$, $\phi_{\mathrm{I}} = 2Ai \sin \alpha a$, while, for $x = -a$, $\phi_{\mathrm{I}} = -2Ai \sin \alpha a$. For the first type of vibration ϕ_{I} has the same value for $+x$ and $-x$

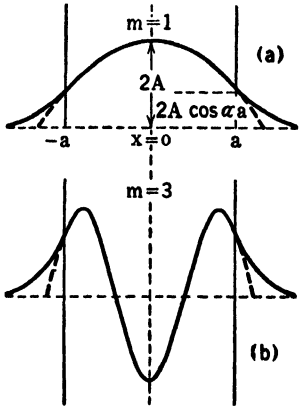


FIG. 13. Symmetrical eigenfunctions for electron in " box."

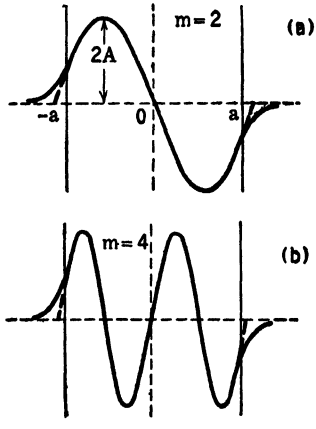


FIG. 14. Antisymmetrical eigenfunctions for electron in " box."

(see Fig. 13) and is known as an *even* function, while for the second type of vibration, ϕ_{I} is antisymmetrical with respect to a change from $+x$ to $-x$ (see Fig. 14) and is known as an *odd* function.[3]

[3] Figure 14 represents, of course, the real function $2A \sin \alpha x$, which is the modulus of the function ϕ_{I} in equation (26). That is, $\phi_{\mathrm{I}}\bar{\phi}_{\mathrm{I}} = (2iA \sin \alpha x)(-2iA \sin \alpha x) = 4 \mid A \mid^2 \sin^2 \alpha x$.

It will be observed that in both these cases there is a definite probability of penetration into regions for which $U_0 > E$. These probabilities are determined by $\phi_{II}\bar{\phi}_{II}$ and $\phi_{III}\bar{\phi}_{III}$ as functions of x, which are real magnitudes in both the symmetrical and antisymmetrical cases.

In the limit, for $\beta = \infty$, that is U_0 infinitely great, $\epsilon^{-2\beta x} = 0$, and there is no penetration into the regions outside the " walls " at $+a$ and $-a$. Under these conditions equation (21) requires that $\tan \alpha a = \infty$ for $\beta/\alpha = \infty$.

That is, αa must be equal to an *odd* multiple of $\pi/2$, and therefore

$$\alpha = \frac{\pi}{2a}(2n + 1); \quad \alpha^2 = \frac{\pi^2}{4a^2}(2n + 1)^2,$$

where $n = 0, 1, 2$, etc.

Since

$$\alpha^2 = \frac{8\pi^2\mu E}{h^2},$$

it follows that

$$E = \left(\frac{2n + 1}{2a}\right)^2 \frac{h^2}{8\mu} = \left(\frac{m}{2a}\right)^2 \frac{h^2}{8\mu}, \tag{27}$$

where $m = 2n + 1$, is an *odd* integer.

Similarly equation (22) requires that $\cot \alpha a = -\infty$ for $\beta = \infty$.

That is, αa must be equal to an *even* multiple of $\pi/2$, and therefore

$$\alpha = \frac{\pi}{2a}(2n),$$

and

$$E = \left(\frac{2n}{2a}\right)^2 \frac{h^2}{8\mu} = \left(\frac{m}{2a}\right)^2 \frac{h^2}{8\mu}, \tag{28}$$

where $n = 1, 2$, etc., and $m = 2n$ is an *even* integer.

Equations (27) and (28) signify that an electron between two potential boundaries, $U_0 > E$, will behave in much the same manner as a stretched string. The electron in the box cannot have a continuously varying set of values for the kinetic energy E, but rather these *energy values will form a series of discrete values*, which for U_0 infinitely large compared with E is defined by the relation

$$E_m = \left(\frac{m}{d}\right)^2 \frac{h^2}{8\mu}, \tag{29}$$

where $m = 1, 2, 3$, etc., and $d = 2a$, denotes the distance between the two boundaries.

These are the *eigenvalues* corresponding to the *eigenfunctions:*

$$\phi_s = 2A \cos \frac{2\pi x}{\lambda}, \tag{30}$$

and

$$\phi_A = 2Ai \sin \frac{2\pi x}{\lambda}, \tag{31}$$

where

$$\lambda = \frac{h}{\sqrt{2\mu E_m}} = \text{de Broglie wave length},$$

$$= \frac{2d}{m} = \frac{4a}{m}. \tag{32}$$

Thus ϕ_s (the even function) corresponds to the odd values of m, while ϕ_A (the odd function) corresponds to the even values of m. These relations are evidently identical with those deduced previously for the wave lengths of the vibrations of a stretched string.

Figures 13 and 14 show the symmetrical (or even) and antisymmetrical (or odd) eigenfunctions, respectively, corresponding to values of m ranging from 1 to 4. It will be observed that the number of nodes (that is, points at which $\phi = 0$) is always equal to $m - 1$.

We shall now attempt to interpret the meaning of ϕ_s or ϕ_A. According to the point of view of quantum mechanics, $\phi\bar{\phi}dx$ is the probability of occurrence of the particle or particles in an element dx at the point x. In the present case

$$\phi_s\bar{\phi}_s = 4A^2 \cos^2 \frac{2\pi x}{\lambda},$$

and

$$\phi_A\bar{\phi}_A = 4A^2 \sin^2 \frac{2\pi x}{\lambda}.$$

These functions are designated as *probability distribution functions*, and from their form it is evident that they are always real and positive.[4] Hence, the value of

$$\int_{-a}^{a} \phi\bar{\phi}dx \text{ is real and positive.}$$

Now, if we wish to describe an experiment in which we know that there is *just one electron in the box*, we must determine the value of A

[4] Plots of these functions are given in Pauling and Wilson, "Introduction to Quantum Mechanics," p. 97.

to be such that the *total probability* of finding the electron in the region $x = -a$ to $x = +a$ is *unity*. (This is, of course, applicable only to the case in which $U_0/E = \infty$.) Hence,

$$\int_{-a}^{a} \phi_s \bar{\phi}_s dx = 1 = 4A^2 \int_{-a}^{a} \cos^2 \frac{2\pi x}{\lambda} \cdot dx$$

$$= 4A^2 \cdot \frac{\lambda}{2\pi} \int_0^{m\pi} \cos^2 \theta d\theta.$$

The limits $x = -a$ and $x = a$ are evidently identical with the limits $x = 0$ and $x = 2a$, and these are identical with the limits $\theta = 2\pi x/\lambda = 0$ and $\theta = 4\pi a/\lambda = m\pi$, as shown by equation (32).

Since

$$\int_0^{\pi} \cos^2 \theta d\theta = \int_0^{\pi} \sin^2 \theta d\theta = \frac{\pi}{2},$$

it follows that

$$1 = 4A^2 \cdot \frac{\lambda}{2\pi} \cdot \frac{m\pi}{2} = 4A^2 \cdot \frac{m\lambda}{4},$$

and

$$2A = \frac{2}{\sqrt{m\lambda}} = \sqrt{\frac{2}{d}}. \tag{33}$$

The same result is obtained from (31) by considering the value of

$$\int_{-a}^{a} \phi_A \bar{\phi}_A dx.$$

This procedure, which consists in determining the value of the coefficient $2A$ which will make the total probability equal to unity, is known as *normalization* of the eigenfunction, and when equation (30) is written in the form

$$\phi_s = \sqrt{\frac{2}{d}} \cdot \cos \frac{2\pi x}{\lambda}, \tag{34}$$

the right-hand side is said to be the *normalized form of the function*, while $\sqrt{d/2}$ is known as the *normalizing factor*. In this equation, λ varies inversely as $\sqrt{E}$, according to equation (32), and both d and λ could be expressed in terms of the eigenvalue E_m, thus:

$$(\phi_s)_m = \frac{2}{\sqrt{mh}} \cdot (2\mu E_m)^{\frac{1}{4}} \cdot \cos \frac{2\pi x \sqrt{2\mu E_m}}{h}, \tag{35}$$

where $m = 1, 3, 5$, etc. The normalized eigenfunctions $(\phi_A)_m$ will have the same coefficients with $m = 2, 4, 6$, etc., and the sine replacing the cosine functions.

As N. F. Mott has pointed out,[5] the eigenvalues E_m are very close together if d is of ordinary dimensions. But as d is decreased toward atomic dimensions, the value of the lowest energy level E_1 increases and the spread between levels also increases.

These conclusions are evident from a consideration of equation (29). For an electron in a box for which $d = 2$ cm.,

$$E_m = \frac{m^2}{4} \cdot \frac{(6.55 \times 10^{-27})^2}{8 \times 9 \times 10^{-28}}$$

$$= m^2 \times 1.49 \times 10^{-27} \text{ erg.}$$

On the other hand, for $d = 2 \times 10^{-8}$ cm.,

$$E_m = \frac{m^2 \times 1.49 \times 10^{-27}}{10^{-16}} = m^2 \times 1.49 \times 10^{-11} \text{ erg}$$

$$= \frac{m^2 \times 1.49 \times 10^{-11}}{1.591 \times 10^{-12}} = 9.15\, m^2 \text{ electron volts}$$

$$= 9.15 \text{ e.v. for } m = 1$$

$$= 36.60 \text{ e.v. for } m = 2, \text{ etc.}$$

(The electron volt is the kinetic energy acquired by an electron when accclerated through a potential difference of 1 volt.)

The energy for the excitation of the first energy level in atomic hydrogen is equivalent to 10.12 electron volts, which is of the same order of magnitude as the energy corresponding to $m = 1$ for an electron in the case $d = 2 \times 10^{-8}$ cm. As mentioned previously an electron bound in a hydrogen atom by the field due to the nucleus resembles an electron between potential barriers. Thus, the solution of this latter problem accounts, qualitatively at least, for the experimental observation that the electron in the hydrogen atom exhibits a series of discrete energy values.

An interesting mathematical deduction from equations (30) and (31) should be mentioned in this connection.

Starting with the function

$$\phi_m = \cos \frac{2\pi x}{\lambda} = \cos \frac{m\pi x}{2a},$$

[5] N. F. Mott, "An Outline of Wave Mechanics," p. 61.

let us consider the integral

$$I = \int_0^{2a} \cos\frac{m_1\pi x}{2a} \cos\frac{m_2\pi x}{2a} \cdot dx,$$

where m_1 is not equal to m_2, $(m_1 \neq m_2)$.

From trigonometrical formulas it follows that

$$I = \frac{1}{2}\int_0^{2a} \cos\left\{(m_1 + m_2)\frac{\pi x}{2a}\right\} dx + \frac{1}{2}\int_0^{2a} \cos\left\{(m_1 - m_2)\frac{\pi x}{2a}\right\} dx$$

$$= \frac{2a}{2\pi(m_1 + m_2)}\int_0^{2a} d \sin\left\{(m_1 + m_2)\frac{\pi x}{2a}\right\} +$$

$$\frac{2a}{2\pi(m_1 - m_2)}\int_0^{2a} d \sin\left\{(m_1 - m_2)\frac{\pi x}{2a}\right\} = 0,$$

since $\sin (m_1 + m_2)\pi = \sin (m_1 - m_2)\pi = \sin 0 = 0$.

The eigenfunctions ϕ_{m_1} and ϕ_{m_2}, where $m_1 \neq m_2$ are said to be *orthogonal* to each other, and we can generalize this result by the statement that the *different eigenfunctions, corresponding to different values of* m, *form an orthogonal set*, that is

$$\int_0^{2a} \phi_{m_1}\bar{\phi}_{m_2} dx = 0 \qquad (m_1 \neq m_2),$$

where the integration is carried out over the whole range in which the functions have physical significance and the bar over the second function indicates that when ϕ_{m_1} is a complex function, the conjugate complex of the second function is to be used in the integration, since only in this manner can the product be made to indicate a real value.

On the other hand, for any normalized eigenfunction, as shown already,

$$\int_0^{2a} \phi_{m_1}\bar{\phi}_{m_1} dx = 1.$$

The trigonometric functions constitute the simplest types of orthogonal functions. Thus, we have

$$\int_0^{2\pi} \cos mx \cdot \cos nx dx = \int_0^{2\pi} \sin mx \cdot \sin nx dx$$

$$= \int_0^{2\pi} \sin mx \cdot \cos nx dx$$

$$= 0 \text{ for } m \neq n, \tag{36}$$

whereas

$$\int_0^{2\pi} \cos^2 nx dx = \int_0^{2\pi} \sin^2 nx dx = \pi. \tag{37}$$

The set of functions

$$\frac{\sin x}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos nx}{\sqrt{\pi}},$$

form a normalized orthogonal (or *orthonormal*) set in the interval 0 to 2π.

In a subsequent chapter it is shown that the different eigenfunctions $\phi_1, \phi_2, \ldots \phi_m \ldots$, corresponding to the eigenvalues $E_1, E_2, \ldots E_m \ldots$ which are obtained as solutions of the S. equation for any form of the function $U(x)$, form an orthogonal set. This is also true for the case in which U and ϕ are each functions of two or more coördinate variables.

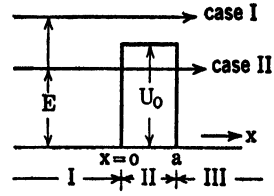


FIG. 15. Illustrating penetration of particle through rectangular potential barrier.

3.3 Transmission of Electrons through Potential Barrier of Finite Extent. The simplest case of a potential barrier of finite thickness is that illustrated in Fig. 15, where $U = 0$ for $x < 0$, $U = U_0$ for the region $x = 0$ to $x = a$, and $U = 0$ for $x > a$.

Case I. $E > U_0$

We now write down the S. equation for each of the three regions and the corresponding general solutions.

(Region I) $$\frac{d^2\phi_I}{dx^2} + \alpha^2\phi_I = 0.$$

The corresponding solution is

$$\phi_I = Ae^{i\alpha x} + Be^{-i\alpha x}.$$

(Region II) $$\frac{d^2\phi_{II}}{dx^2} + \beta^2\phi_{II} = 0,$$

where $$\beta^2 = \frac{8\pi^2\mu(E - U_0)}{h^2}.$$

Solution is $$\phi_{II} = Ce^{i\beta x} + De^{-i\beta x}.$$

(Region III) $$\frac{d^2\phi_{III}}{dx^2} + \alpha^2\phi_{III} = 0.$$

Solution is

$$\phi_{III} = F \cdot \epsilon^{i\alpha x},$$

since the beam of electrons is transmitted from region I through II and III, and there are no electrons moving from right to left in region III.

Postulating as in the previous problems that, at $x = 0$,

$$\phi_I = \phi_{II},$$

$$\frac{d\phi_I}{dx} = \frac{d\phi_{II}}{dx},$$

and at $x = a$,

$$\phi_{II} = \phi_{III},$$

$$\frac{d\phi_{II}}{dx} = \frac{d\phi_{III}}{dx},$$

we derive the relations

$$\left.\begin{aligned}
\frac{C}{F} &= \frac{1}{2}\left(1 + \frac{\alpha}{\beta}\right)\epsilon^{ia(\alpha-\beta)} \\
\frac{D}{F} &= \frac{1}{2}\left(1 - \frac{\alpha}{\beta}\right)\epsilon^{ia(\alpha+\beta)} \\
\frac{A}{F} &= \frac{1}{2}\left[\left(1 + \frac{\beta}{\alpha}\right)\frac{C}{F} + \left(1 - \frac{\beta}{\alpha}\right)\frac{D}{F}\right] \\
\frac{B}{F} &= \frac{1}{2}\left[\left(1 - \frac{\beta}{\alpha}\right)\frac{C}{F} + \left(1 + \frac{\beta}{\alpha}\right)\frac{D}{F}\right]
\end{aligned}\right\}. \tag{38}$$

The number of electrons incident per unit time on the barrier at $x = 0$ is

$$I_i = |A|^2 v_0,$$

where $|A|^2$ denotes the concentration per unit length of electrons traveling in the direction of increasing values of x, and v_0 is the velocity, defined by the relation $\mu v_0 = \sqrt{2\mu E} = h\alpha/(2\pi)$. (As mentioned previously, the notation $|A|^2$ is used to designate the numerical value of $A\bar{A}$.)

This incident beam is partly reflected at $x = 0$, and the rate at which electrons are reflected is given by

$$I_r = |B|^2 v_0.$$

The rate at which the electrons are transmitted into the barrier is given by

$$I_t = |C|^2 v,$$

where

$$\mu v = \sqrt{2\mu(E - U_0)} = \frac{h\beta}{2\pi}.$$

This transmitted beam is partly reflected at $x = a$, and the rate of reflection is

$$I_r' = |D|^2 v,$$

while that of transmission into the region $x > a$ is

$$I_t' = |F|^2 v_0.$$

Thus, the net *transmission coefficient* for electrons incident on the barrier at $x = 0$ is given by

$$T = \frac{|F|^2}{|A|^2}. \tag{39}$$

Since only ratios of constants occur in (39) we will put F equal to *unity*, and hence the transmission coefficient for electrons through the barrier is given by $|1/A|^2$. From relations (38) it follows that

$$A = \frac{1}{4}\left[\left(1 + \frac{\beta}{\alpha}\right)\left(1 + \frac{\alpha}{\beta}\right)\epsilon^{ia(\alpha-\beta)} + \left(1 - \frac{\beta}{\alpha}\right)\left(1 - \frac{\alpha}{\beta}\right)\epsilon^{ia(\alpha+\beta)}\right]$$

$$|A|^2 = A\bar{A} = \frac{1}{16}\left[\left(1 + \frac{\beta}{\alpha}\right)^2\left(1 + \frac{\alpha}{\beta}\right)^2 + \left(1 - \frac{\beta}{\alpha}\right)^2\left(1 - \frac{\alpha}{\beta}\right)^2 + \left(1 - \frac{\alpha^2}{\beta^2}\right)\left(1 - \frac{\beta^2}{\alpha^2}\right)(\epsilon^{2ia\beta} + \epsilon^{-2ia\beta})\right].$$

In this equation we may replace the exponential term by $\cos 2a\beta = \cos(4\pi a/\lambda)$ where λ is the de Broglie wave length inside the barrier, corresponding to the momentum $\mu v = h\beta/2\pi$.

As E is increased more and more with respect to U_0, β asymptotically approaches α, and $|A|^2 = 1$ in the limit. That is, there is complete transmission of electrons through the barrier.

Case II. $E < U_0$

According to classical mechanics, the particles incident at the barrier are completely reflected. Quantum mechanics leads to a different solution, which is of extremely great physical significance.

The solutions of the S. equation for this case will differ from those given for case I in this respect, that β corresponds to an imaginary de Broglie wave length, and hence we have the three relations:

$$(x < 0) \qquad \phi_{\text{I}} = A\epsilon^{i\alpha x} + B\epsilon^{-i\alpha x},$$

$$(0 < x < a) \qquad \phi_{\text{II}} = C\epsilon^{\beta x} + D\epsilon^{-\beta x},$$

$$(x > a) \qquad \phi_{\text{III}} = \epsilon^{i\alpha x},$$

where the coefficient in the last relation is put equal to unity for the same reason as in case I.

Putting in the conditions valid at $x = 0$ and $x = a$, we obtain the following values for the four constants:

$$A = \frac{\epsilon^{i\alpha a}}{4}\left[2(\epsilon^{\beta a} + \epsilon^{-\beta a}) + i\left(\frac{\beta}{\alpha} - \frac{\alpha}{\beta}\right)(\epsilon^{\beta a} - \epsilon^{-\beta a})\right]; \tag{40}$$

$$B = -\frac{i\epsilon^{i\alpha a}}{4}\left(\frac{\beta}{\alpha} + \frac{\alpha}{\beta}\right)(\epsilon^{\beta a} - \epsilon^{-\beta a}); \tag{41}$$

$$C = \frac{1}{2}\left(1 + \frac{i\alpha}{\beta}\right)\epsilon^{i\alpha a - \beta a};$$

$$D = \frac{1}{2}\left(1 - \frac{i\alpha}{\beta}\right)\epsilon^{i\alpha a + \beta a}.$$

In these equations, $A\bar{A}v_0$ represents the " current " for the incident particles, $B\bar{B}v_0$ that for the reflected particles, and v_0 that for the particles transmitted through the barrier at $x = a$.

It can be shown very readily from equations (41) and (40) that

$$A\bar{A} - B\bar{B} = 1,$$

while the probability that a particle coming up to the boundary at $x = 0$ shall " tunnel " through the barrier is

$$P\left(\frac{E}{U}\right) = \frac{1}{A\bar{A}},$$

where $P(E/U)$ designates that this probability is a function of the ratio E/U.

From equation (40), it follows that

$$A = \epsilon^{i\alpha a}\left\{\cosh \beta a + \frac{i}{2}\left(\frac{\beta}{\alpha} - \frac{\alpha}{\beta}\right)\sinh \beta a\right\},$$

while

$$\bar{A} = \epsilon^{-ia\alpha}\left\{\cosh \beta a - \frac{i}{2}\left(\frac{\beta}{\alpha} - \frac{\alpha}{\beta}\right)\sinh \beta a\right\}.$$

Hence,

$$A\bar{A} = \cosh^2 \beta a + \frac{1}{4}\left(\frac{\beta}{\alpha} - \frac{\alpha}{\beta}\right)^2 \sinh^2 \beta a.$$

For U_0 very much greater than E, β/α is very large, and $\cosh \beta a = \sinh \beta a$. Hence we obtain the approximation

$$A\bar{A} = \frac{1}{4} \cdot \frac{\beta^2}{\alpha^2} \sinh^2 \beta a,$$

$$= \frac{1}{16} \cdot \frac{\beta^2}{\alpha^2} \cdot \epsilon^{2\beta a},$$

and

$$P\left(\frac{E}{U}\right) = \frac{16\alpha^2}{\beta^2} \cdot \epsilon^{-2\beta a}$$

$$= \frac{16\dfrac{E}{U_0} \cdot \epsilon^{-2\beta a}}{1 - \dfrac{E}{U_0}}. \tag{42}$$

Except for values of E/U_0 which are near 1 or 0, the exponential term is the important factor, and the equation indicates that except for values of E, small compared to those of U_0, there is a finite and measurable probability that a particle will get through a potential barrier if the latter has a width a which is of the same order of magnitude as

$$\frac{1}{\beta} = \frac{h}{2\pi\sqrt{2\mu(U_0 - E)}}.$$

Thus, let us consider an electron for which, if $(U_0 - E)$ is expressed in electron volts V,

$$\frac{1}{\beta} = \frac{3.89 \times 10^{-8}}{\sqrt{V}} \text{ cm.}$$

Assume $E = 1$, and $U_0 = 10$ in electron volts, then

$$\frac{1}{\beta} = 1.30 \times 10^{-8} \text{ cm.},$$

and we obtain the following values of P for different values of a:

$a \times 10^8$	P
1.30	0.248
1.95	0.089
2.60	0.035
6.50	10^{-4} approx.

Thus P decreases rapidly as a is increased, and, for values of $a > 10$ Å, P becomes infinitesimally small; but for atomic dimensions, the probability is evidently fairly large.

This phenomenon of the penetration of particles through potential barriers, or the "tunneling effect," as it has been designated, is one of the most important deductions contributed by the new quantum mechanics. R. W. Gurney and E. U. Condon[6] have described the significance of this conclusion as follows:

In classical mechanics the orbit of a particle is entirely confined to those points in space at which its potential energy is less than its total energy. This is not true in quantum mechanics. Classically if a particle be moving in a basin of low potential energy and have not as much total energy as the maximum of potential energy surrounding the basin, it must *certainly* remain there for all time, unless it acquires the deficiency in energy somehow. But in quantum mechanics most statements of certainty are replaced by statements of probability. And the above statement must now be altered to read ". . . it may remain there for a long time but as time goes on, the probability that it has escaped, even without change in its total energy, increases towards unity . . . ."

For instance, let us consider a particle of mass μ and the total energy E, moving in a range where the potential energy function $U(x)$ is of the form shown in Fig. 16. If E is less than the maximum value for the height of the "hill" between the two "valleys," then according to classical mechanics there would be two different types of motion possible for the particle, each of which is confined to one of the regions I or II. This motion would be of the nature of a vibration, with different frequencies in each region.

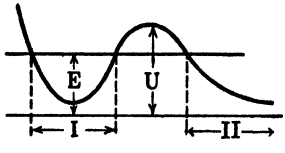


FIG. 16. Illustrating penetration of particle through barrier of arbitrary shape.

But from the new point of view, "There is no longer a definite correlation between simultaneous value of position and momentum as implied by the equation

$$\frac{p^2}{2\mu} + U(x) = E."$$

Instead of regarding the particle as moving with definite velocity at

[6] R. W. Gurney and E. U. Condon, *Phys. Rev.*, **33**, 127 (1929).

any given point, we forget about the particle and consider, instead, the properties of the associated wave motion and its " amplitude " function ϕ which is a function of both x and E. The square of this function, or rather $\phi\bar{\phi}dx$, is interpreted as giving the probability that the particle lies between x and $x + dx$, when it is in the state of energy E. " This," as Gurney and Condon point out, " is really the ground for requiring that ϕ remain finite. For an energy level, such that $\phi(E, x)$ does not remain finite as $x \rightarrow \pm \infty$, the probability that it is not at infinity is vanishingly small, and therefore these states do not exist physically. Adopting the probability interpretation of $\phi(E, x)$ one has at once the result that there is a finite probability of being outside the range of classical motion of that energy."

The point of view thus emphasized in the remarks from the paper of Gurney and Condon and illustrated by the problems which have been solved in the previous section has received a number of important applications. As a preliminary to the consideration of these let us consider the following problem.

For the electrons confined between two semi-infinite barriers, it was shown that, for values of E not too small compared with those of U, there is a definite probability of penetration into the regions $U > E$. If now these latter regions are decreased in extent until they become of atomic dimensions, the probability becomes measurable that the electrons which have entered the barriers will also pass completely through them, and this probability will assume a definite value for each discrete energy state E_m of the electrons in the " box." To indicate this we may denote the probability by $P(E_m)$.

Without attempting to solve this problem in detail, it is readily seen that, as in Fig. 12, there will exist a series of discrete energy states, designated by $E_1, E_2 \ldots E_m$, and that the corresponding eigenfunctions $\phi(E_m, x)$ will be of the nature of sine and cosine functions of the angle αx, where

$$\alpha = \frac{2\pi}{h}\sqrt{2\mu E_m}.$$

Furthermore, it is evident that, if b denote the width of the region of atomic dimensions, for which $U > E$, the probability of penetration will be given, according to equation (42), by a relation of the form

$$P(E_m) = F\left(\frac{E_m}{U}\right) \cdot \epsilon^{-2\beta b}, \tag{43}$$

where

$$2\beta b = \frac{4\pi b}{h}\sqrt{2\mu(U - E)}, \tag{44}$$

and F is a function of the ratio E_m/U. Except for values of this ratio approaching zero or unity, the controlling factor governing the value of P will be the exponential term.

For those cases in which the value of U in the region $U > E$ is a function of x, denoted by $U(x)$, the value of the exponent $2\beta b$ is given, approximately, by the expression

$$2\beta b = \frac{4\pi}{h}\int \sqrt{2\mu(U - E)} \cdot dx, \tag{45}$$

where the integration extends across the barrier between the two limits at which $U(x) - E = 0$. It is readily seen that for $U(x) = U_0$ in the range $x = \pm a$ to $x = \pm(a + b)$, and $U(x) = 0$ in the other regions, this integral gives the result stated in (44).

3.4 Application to Radioactive Disintegration. As first shown by R. W. Gurney and E. U. Condon[7] and, independently, by G. Gamow,[8] the theory of the penetration of particles through potential barriers of atomic dimensions gives a very satisfactory interpretation of the phenomena of ejection of alpha and beta particles from nuclei of the radioactive elements.

As is well known this emission of charged corpuscles occurs spontaneously and leads to the formation of a new atomic species. The rate of disintegration of any nuclear type is given by

$$-dN = N\gamma dt,$$

where N is the number of nuclei which are unaltered at the end of any given period t, and γ is known as the decay constant. This leads to the exponential law of decay

$$N(t) = N_0\epsilon^{-\gamma t},$$

where N_0 is the number present at $t = 0$. It is customary to put $\gamma = 1/\tau$ where τ is known as the mean " life " of the nucleus.

For any given nuclear type, the alpha particles emitted are all of the same velocity v, or at the most possess two or three discrete values of v. On the other hand, the velocities of the beta particles form a continuous range of values. For the emission of alpha particles we have the empirical relation, known as the Geiger-Nuttall law, according to which

$$\log \gamma = A + B \log v,$$

where A and B are constants. That is, the life of a nuclear structure is shorter, the higher the velocity of the emitted alpha particles.

[7] R. W. Gurney and E. U. Condon, *loc. cit.*

[8] G. Gamow, *Z. Physik.*, **51**, 204 (1928).

In order to account for these observations we have to consider the nature of the potential energy function for an alpha particle in the neighborhood of the nucleus. From experiments on the scattering of alpha particles it is known that the inverse square law of repulsive forces applies to within a distance of about 10^{-12} cm. from the nucleus. If we denote the charge on the nucleus by $+Ze$ and that on the alpha particle by $+2e$, the force, as a function of the distance r, is given by $2Ze^2/r^2$, and, therefore, as will be shown in Chapter IV,

$$U = \int_r^\infty \frac{2Ze^2}{r^2} \cdot dr = \frac{2Ze^2}{r}.$$

This is indicated by the hyperbolic portion of the curve in Fig. 17.

However, in order to account for the fact that alpha particles remain within the nuclear structure for a definite period of the order τ, it is necessary to assume that very close to the nucleus the field changes from

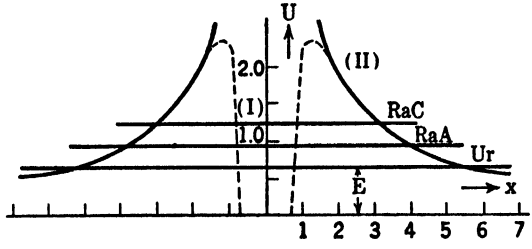


FIG. 17. Potential energy curve for alpha particle in nucleus of radioactive atom.

repulsive to attractive. Therefore, the function U will be represented very close to the nucleus by the dotted portion in the figure. That is, the nucleus is surrounded by a potential barrier of maximum height U_0, and if the energy of the emitted alpha particle be represented by $E < U_0$, then according to the arguments which have been given in the previous section, there exists a definite probability for the penetration of the particles through the barrier. The following remarks on the calculation of this probability follow closely the presentation of Gurney and Condon.

Evidently the probability of emission of an alpha particle will be equal to the reciprocal of the mean time τ, during which a particle remains in the region (I) close to the nucleus before " leaking through " to the outer region (II) where it experiences a repulsive field. For a particle of kinetic energy E, the velocity for $U = 0$ is $v = \sqrt{2E/\mu}$. Hence, the amount of time spent by the particle in unit distance for x very large is $\sqrt{\mu/(2E)}$, and the time spent in a range of length a is therefore $a\sqrt{\mu/(2E)}$.

Now according to quantum mechanics, as shown in equation (43), the probability of penetration is of the order

$$P \approx \epsilon^{-2\beta b},$$

where P is defined as the ratio between the probability of occurrence of the particle in unit length of the region II and the probability of occurrence in unit length of the region I. "Therefore since the motion is aperiodic and the particle escaping from the range I will in the mean only go through unit length of II once, the time τ which must be spent in range I before getting through to range II is of the order

$$\tau \approx a\sqrt{\frac{\mu}{2E}} \cdot \epsilon^{2\beta b},$$

where $2\beta b$ is defined by equation (45) and a is of the order of the breadth of range I."

Hence, the decay constant is given by the relation

$$\gamma = \frac{1}{\tau} \approx \frac{1}{a}\sqrt{\frac{2E}{\mu}} \cdot \epsilon^{-2\beta b}, \tag{46}$$

where the value of $2\beta b$ depends upon *the distance between the maximum value of* U *and the horizontal line corresponding to the particular value of* E *for the emitted alpha particles.* It is evident from this conclusion "that if the size of the potential barrier be increased by a small factor the probability of escape may be decreased more than a million fold."

To test this theory, Gurney and Condon have calculated the relative lives of Ur, RaA, and RaC′ from the observed values for the kinetic energies of the emitted alpha particles, which are 6.5×10^{-6}, 9.55×10^{-6}, and 12.2×10^{-6} erg, respectively. These values of E are indicated by the horizontal lines in Fig. 17, while the potential energy function for an alpha particle with respect to the nucleus is indicated by the curve. From the areas of the portions between this curve and the horizontal lines, values of the exponential factor in equation (46) were calculated, and these were then used to calculate relative values of γ, the decay constant. The agreement between these and the observed relative values was found to be very satisfactory. Thus, "it is the energy of the emitted alpha particle which determines its own rate of decay," as had been concluded previously from the Geiger-Nuttall relation.

3.5 Emission of Electrons in Presence of Intense Fields. It has been shown by a number of investigators that, with extremely high

electric fields, electron emission from cathodes occurs even at room temperature. This is known as the " cold-cathode " or " auto-electronic " effect. An explanation of the phenomenon on the basis of the quantum mechanics has been given by J. R. Oppenheimer[9] and by R. H. Fowler and L. Nordheim.[10]

As mentioned previously electron emission occurs from a cathode when the kinetic energy is equal to or exceeds the work function. With ordinary electric fields the potential energy function for the electrons resembles that shown in Fig. 9, and the kinetic energy becomes sufficiently high to overcome the barrier only at higher temperatures. But in the presence of extremely high electric fields, even at ordinary temperatures, the potential energy function is represented by the plot in Fig. 18 where

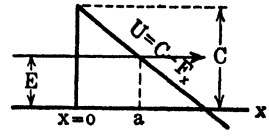


FIG. 18. Illustrating "cold-cathode" emission of electrons.

$$U(x) = 0 \text{ for } x < 0$$

$$U(x) = C - Fx \text{ for } x \geqslant 0$$

and C denotes the work function, that is, the value of the potential energy at $x = 0$, while F is the field strength (corresponding to the voltage gradient).

The S. equation which is to be solved in this case is of the form

$$\frac{d^2\phi}{dx^2} + \frac{8\pi^2\mu}{h^2}(E - C + Fx)\phi = 0, \tag{47}$$

where E, the kinetic energy of the electrons, is less than C. This is a differential equation which differs from those solved in the previous sections in that the coefficient in the second term on the left-hand side is not a constant, but a linear function of x. However, it is evident from the considerations advanced already that if the distance a, indicated in Fig. 18 as the width of the barrier, is of atomic dimensions, there will exist a probability of significant magnitude for the penetration of electrons through the barrier.

From equation (45) it follows that this probability will be of the form

$$P(E) = K\epsilon^{-2\beta b}, \tag{48}$$

[9] J. R. Oppenheimer, *Phys. Rev.*, **31**, 66 (1928).

[10] R. H. Fowler and L. Nordheim, *Proc. Roy. Soc.*, **A119**, 173 (1928); L. Nordheim, *Physik. Z.*, **30**, 177 (1929). The investigations on this topic have been reviewed by S. Dushman, *Rev. Modern Phys.*, **2**, 381 (1930).

where K is a function of E/C, and

$$2\beta b = \frac{4\pi}{h}\int_0^a \sqrt{2\mu(U-E)}\cdot dx$$

$$= \frac{4\pi}{h}\int_0^a \sqrt{2\mu(C-Fx-E)}\cdot dx.$$

Let $C - E = A$. Then

$$d(A-Fx)^{\frac{3}{2}} = -\left(\frac{3}{2}\right)\cdot F(A-Fx)^{\frac{1}{2}}dx,$$

and

$$(A-Fx)^{\frac{3}{2}}\Big]_a^0 = \frac{3F}{2}\int_0^a (A-Fx)^{\frac{1}{2}}dx.$$

For $x = a, Fa = C - E = A$. Therefore,

$$\frac{4\pi\sqrt{2\mu}}{h}\cdot\frac{2}{3F}(C-E)^{\frac{3}{2}} = \frac{4\pi}{h}\int_0^a \sqrt{2\mu(C-Fx-E)}\cdot dx,$$

and

$$2\beta b = \frac{8\pi}{h}\cdot\frac{(2\mu)^{\frac{1}{2}}(C-E)^{\frac{3}{2}}}{3F}. \tag{49}$$

It will be observed that the constant

$$\frac{1}{F} = \frac{a}{C-E}$$

takes the place of a, the extent of the barrier, which is the physical significance of b in the exponential term $\epsilon^{-2\beta b}$.

According to Fowler and Nordheim, the coefficient K in (48) is given " with sufficient accuracy " by the relation

$$K = 4\left(\frac{E}{C}\right)^{\frac{1}{2}}\left(1-\frac{E}{C}\right)^{\frac{1}{2}},$$

so that the resulting expression for $P(E)$ in (48) is very similar to that given in (42) for the probability of penetration where $U > E$ is constant over the extent of the barrier.

From these results, it is possible to derive a relation between the electron emission and the field strength which is in agreement with the experimental observations.

The theory of the penetration of electrons through potential barriers has also been applied by Fowler and Nordheim to account for the observations on the increase in emission from tungsten when covered by a monatomic film of a more electropositive metal such as thorium.[11] In this case the potential energy function is of the form shown in Fig. 19, where W_b denotes the work function for the pure metallic surface, and W_a denotes the work function when a monatomic film of thickness l covers the surface. The kinetic energy of the electrons in the metal is denoted by W, and it is to be expected that for values of W between those of W_a and W_b there will exist a probability of transmission of electrons through the film, which is of the form

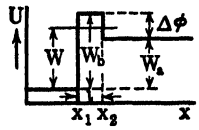


FIG. 19. Potential distribution at surface of metal covered with monatomic film of more electropositive element.

$$P = K \cdot \epsilon^{-2\beta l},$$

where

$$\beta = \frac{2\pi}{h}\sqrt{2\mu(W_b - W_a)}.$$

According to L. Nordheim and R. H. Fowler,[12]

$$K = \frac{8(\pi k T W_a)^{\frac{1}{2}}}{W_b}$$

where k = Boltzmann constant = 1.37×10^{-16} erg/deg., and T = absolute temperature.

In this connection the reader will find it of interest to study the paper by C. Eckart[13] in which the potential function $V(x)$ is represented by an analytic function of the form

$$V(x) = -\frac{A\epsilon^{-\alpha x}}{1 - \epsilon^{-\alpha x}} - \frac{B\epsilon^{-\alpha x}}{(1 - \epsilon^{-\alpha x})^2},$$

where $\alpha = 2\pi/d$, and the "barrier" is represented by a curve which extends from $x = -d$ to $x = +d$, with the values $V = 0$ at the first point and V = constant, for $x \geq d$.

The resulting S. equation is of the hypergeometric type and the solutions may therefore be expressed in terms of the hypergeometric series.[14]

[11] See Dushman, *loc. cit.*, for a detailed discussion of the observations and their interpretation on the basis of quantum mechanics.

[12] R. H. Fowler, *Proc. Roy. Soc.*, **A122**, 36 (1929).

[13] C. Eckart, *Phys. Rev.*, **35**, 1303 (1930).

[14] These series are discussed in most of the mathematical treatises mentioned at the end of Chapter II.

COLLATERAL READING

1. PAULING L., and WILSON, E. B., Jr., "Introduction to Quantum Mechanics," pp. 90–100, contains a discussion of the three-dimensional case of the free particle and also of the electron in a box.
2. MOTT, N. F., "An Outline of Wave Mechanics," pp. 59–64, gives a well-worthwhile presentation of "normal modes of a de Broglie wave."
3. FRENKEL, J., "Wave Mechanics," Vol. I. The treatment of potential barrier problems given by this writer is extremely clear.
4. CONDON, E. U., "Quantum Mechanics of Collision Processes," *Rev. Modern Phys.*, **3**, 43 (1931). The problems of potential barriers in one dimension are discussed on pp. 51–64.
5. GURNEY, R. W., and CONDON, E. U., *Phys. Rev.*, **33**, 127 (1929). This original publication is extremely well worth a thorough study for the clarity of presentation of a new and fundamental idea in quantum mechanics.

CHAPTER IV

CLASSICAL THEORY OF ATOMIC DYNAMICS[1]

4.1 Relation between Quantum Mechanics and Classical Dynamics. Mention has been made in previous chapters of the Hamiltonian form of expression for the total energy of a system. The formulation of the S. equation for any given system involves, as a first step, the description of the system in the Hamiltonian form, and it is therefore necessary to know how to deduce the latter for any type of system involving two or more particles and two or more coördinate variables. Furthermore, we shall find that a number of the concepts, such as the Principle of Least Action, which are of fundamental significance in classical dynamics, have received an extended application in the new quantum mechanics. The Hamiltonian canonical equations find their analog in the Poisson bracket type of relations which Dirac has utilized in his technic of operators, and in numerous other applications of the new quantum mechanics we find illustrations of the extension or modification of conceptions which were developed in connection with classical dynamics.

It is for this reason, and because we believe that the new ideas cannot be grasped fully except in the light of the older developments, that this chapter has been written. The important fact that must be realized is that the new quantum mechanics represents a type of evolution from the Newtonian dynamics — a development, which is a logical consequence of the Principle of Indeterminacy, of the fact that whereas the classical methods were intended to deal with macroscopic phenomena, the newer modifications are needed in order to deal adequately with atomic phenomena.

4.2 Kinetic and Potential Energy. The laws of classical dynamics are based upon Newton's three laws of motion, of which the second is the most important since it gives us a measure of force.

Let x denote the distance traversed by the particle of mass μ at time t, under action of a force F, and let v denote the velocity at this instant.

Then, according to Newton's second law, the force is defined as the

[1] This chapter is based (with modifications at different places) on the author's discussion of this topic in Taylor's "Treatise on Physical Chemistry," pp. 1264–1280.

rate of change of momentum, that is

$$F = \frac{d}{dt}(\mu v).$$

In non-relativity mechanics, μ is regarded as constant (independent of v), and therefore,

$$F = \mu \cdot \frac{dv}{dt} = \mu \cdot \frac{d^2x}{dt^2} \equiv \mu\ddot{x}, \tag{1}$$

where $\ddot{x}$ is known as the acceleration.

We can write this equation in the form

$$Fdx = \mu \cdot \frac{dv}{dt} \cdot vdt = \frac{1}{2}\mu d(v^2).$$

If v_0 and v_1 denote the values of the velocity at x_0 and x_1, respectively,

$$T_1 - T_0 = \tfrac{1}{2}\mu v_1^2 - \tfrac{1}{2}\mu v_0^2 = \int_{x_0}^{x_1} Fdx. \tag{2}$$

The integral on the right-hand side denotes the *work which is done upon the particle by the force* F, and the left-hand side denotes the *change in kinetic energy* T. Thus, equation (2) signifies that the work done upon the particle is equal to the increase in kinetic energy of the latter.

In general the system under consideration will consist of n particles moving under the action of impressed forces in three-dimensional space. Applying equation (1) to each of these particles we obtain the n sets of relations

$$\left.\begin{aligned} \mu_i\ddot{x}_i - X_i &= 0 \\ \mu_i\ddot{y}_i - Y_i &= 0 \\ \mu_i\ddot{z}_i - Z_i &= 0 \end{aligned}\right\}, \tag{3}$$

where μ_i is the mass of the ith particle, and X_i, Y_i, and Z_i are the three components of the force acting on this particle.

The kinetic energy of the system is given by the relation

$$T = \sum_{i=1}^{n} \tfrac{1}{2}\mu_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2), \tag{4}$$

where $\dot{x}_i = dx_i/dt$, and the summation is taken over the n particles.

It follows that

$$\frac{\partial T}{\partial \dot{x}_i} = \mu_i\dot{x}_i,$$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}_i}\right) = \mu_i\ddot{x}_i = X_i, \tag{5}$$

and similar equations apply to $\ddot{y}_i$ and $\ddot{z}_i$.

In the case of conservative systems (which is the only type that will be considered in this chapter), it is possible to determine a function of the $3n$ coördinates $x_1\ y_1\ z_1\ \ldots\ x_n\ y_n\ z_n$, designated by V, which possess the property that the force components are equal to the negative partial derivatives of V with respect to the corresponding coördinates. That is

$$\left.\begin{aligned} X_i &= -\frac{\partial V}{\partial x_i} \\ Y_i &= -\frac{\partial V}{\partial y_i} \\ Z_i &= -\frac{\partial V}{\partial z_i} \end{aligned}\right\}, \tag{6}$$

where $i = 1, 2, \ldots n$. We can write this set of relations in the form

$$\sum_{i=1}^{n}(X_i dx_i + Y_i dy_i + Z_i dz_i) = -dV, \tag{7}$$

thus indicating that the left-hand side of the last equation is an *exact differential.*

From equations (6) and (5) it follows that we can express Newton's second law in the form of the following equations, of which there are $3n$ for the n particles:

$$\left.\begin{aligned} \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}_i}\right) + \frac{\partial V}{\partial x_i} &= 0 \\ \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{y}_i}\right) + \frac{\partial V}{\partial y_i} &= 0 \\ \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{z}_i}\right) + \frac{\partial V}{\partial z_i} &= 0 \end{aligned}\right\} \tag{8}$$

Now let us consider the set of equations (3). Multiplying these equations respectively by the arbitrary small displacements[2] δx_i, δy_i, and δz_i, and taking the sum of the displacements for all the particles of the system, we obtain the relation

$$\sum_i \left[(\mu_i \ddot{x}_i - X_i)\delta x_i + (\mu_i \ddot{y}_i - Y_i)\delta y_i + (\mu_i \ddot{z}_i - Z_i)\delta z_i \right] = 0. \tag{9}$$

[2] The symbol δ is used to indicate an arbitrary infinitesimal variation.

This relation is the analytical statement of what has been designated d'Alembert's Principle. This was first published in 1743 and has been regarded as the basic equation of dynamics. In fact, " Lagrange made this principle the basis of the entire subject of dynamics."[3]

In equation (9) let us introduce the substitutions $\delta x_i = \dot{x}_i dt$; $\delta y_i = \dot{y}_i dt$; $\delta z_i = \dot{z}_i dt$. This yields the relation

$$\sum_i \left[\mu_i(\ddot{x}_i\dot{x}_i + \ddot{y}_i\dot{y}_i + \ddot{z}_i\dot{z}_i)dt - (X_i\dot{x}_i + Y_i\dot{y}_i + Z_i\dot{z}_i)dt \right] = 0. \quad (10)$$

But

$$\mu_i\ddot{x}_i\dot{x}_i = \frac{1}{2}\frac{d}{dt}(\mu_i\dot{x}_i^2),$$

and

$$X_i\dot{x}_i dt = X_i dx.$$

Hence equation (10) becomes

$$\frac{dT}{dt} \cdot dt + dV = 0. \quad (11)$$

Integrating between the limits $t = 0$ and $t = t$, this last relation becomes

$$T_t - T_0 + V_t - V_0 = 0,$$

that is,

$$T + V = E, \quad (12)$$

where E is a constant which defines the *total energy* of the system.

As an application of equation (10) or (12), let us consider the motion of a single particle along the x-axis under the action of a force F which is a function of x. Hence, we can determine V as a function of x. Given the total energy E, we obtain the relations

$$\frac{\mu v^2}{2} = E - V,$$

$$v = \frac{dx}{dt} = \sqrt{\frac{2(E - V)}{\mu}}, \quad (13)$$

and hence

$$t - t_0 = \int_{x_0}^{x} \frac{dx}{\sqrt{\frac{2(E - V)}{\mu}}}. \quad (14)$$

[3] A. G. Webster, "Dynamics of a Particle."

For instance, for a body moving under gravity,[4] $V = \mu g x$, where g is the gravitational constant. Therefore equation (14) becomes

$$t - t_0 = \int \frac{dx}{\sqrt{\frac{2E}{\mu} - 2gx}}. \tag{i}$$

Let $z = 2E/\mu - 2gx$.

Then $dz = -2g\,dx$, and

$$t - t_0 = -\frac{1}{2g}\int \frac{dz}{\sqrt{z}} = -\frac{\sqrt{z}}{g} = -\frac{1}{g}\sqrt{\frac{2E}{\mu} - 2gx}. \tag{ii}$$

For $x = E/(\mu g)$, $t = t_0$, which shows that this is the highest point from which the body begins to fall. Let $E/(\mu g) = x_0$. Then we can write (ii) in the form

$$t - t_0 = -\sqrt{\frac{2}{g}(x_0 - x)}$$

that is,

$$x - x_0 = -\tfrac{1}{2}g(t - t_0)^2. \tag{iii}$$

In this connection it is essential to make some remarks regarding the signs of E and V in atomic dynamics. Obviously the kinetic energy T is always positive. With regard to the sign of V the case is different. Since measurements yield only the difference in potential energy with respect to some point in space, which is assumed to be the zero of potential, it is necessary in dealing with atomic systems to define arbitrarily the state for which $V = 0$.

Now let us consider an electron (charge $= -e$) at an infinite distance from a positive charge of magnitude Ne. As the electron approaches the latter the force of attraction increases in accordance with Coulomb's law, and the kinetic energy T increases. Since E is constant, V must decrease as r, the distance of separation, decreases.

If now we assume that for $r = \infty$, $V = 0$, then $V = V(r)$ must become *more and more negative as* r *decreases*. That is,

$$V = -\int_{\infty}^{r} F\,dr = -\int_{\infty}^{r}\left(-\frac{Ne^2}{r^2}\right)dr = -\frac{Ne^2}{r}. \tag{15}$$

This shows that V will always be negative if the *negative* sign is used in the relation

$$F = -\frac{Ne^2}{r^2}.$$

[4] Slater and Frank, "Introduction to Theoretical Physics," p. 42.

That is, *for attractive forces the negative sign must be used*, and evidently a positive sign will indicate repulsion.

Since V is negative the sign of E depends upon the absolute magnitudes of T and V. But in atomic systems, consisting of a positive nucleus and one or more electrons, energy must be absorbed by the system in order to remove an electron, that is, work must be done on the system. Hence, E as well as V must be negative. The same conclusion obviously applies to any other system in which attractive forces exist between the constituents. On the other hand, in the case of repulsive forces between the particles of a system E must be positive. The importance of these considerations will become evident from illustrations which will be given both in this and succeeding chapters.

4.3 Generalized Coördinates. Instead of specifying the position of a particle by rectangular coördinates, it is more convenient in many problems to use some other system of coördinates. For instance, the position of the point P may be specified, as shown in Fig. 20, by the distance r and the angles α, β, γ which OP makes with the three rectangular axes. Since

$$\left.\begin{aligned} x &= r\cos\alpha \\ y &= r\cos\beta \\ z &= r\cos\gamma \end{aligned}\right\}, \tag{16}$$

it follows that

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1. \tag{17}$$

Therefore the three angles are not independent, and the position of P can be specified by r and any two of the three angles.

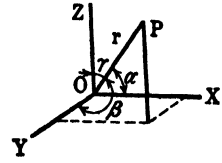


FIG. 20. Illustrating equation (16).

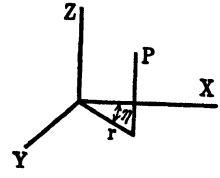


FIG. 21. Illustrating cylindrical coördinates.

Other types of coördinate systems are often used in dynamical problems because of greater convenience. In *cylindrical coördinates* the position of a particle is specified by the coördinates z, r, and η, as shown in Fig. 21. Evidently

$$x = r\cos\eta \tag{18a}$$

$$y = r\sin\eta \tag{18b}$$

$$z = z. \tag{18c}$$

The most convenient system for problems of central orbits, such as those dealing with the motion of an electron about a nucleus, is that of *spherical coördinates*. These are indicated in Fig. 22*a*, where $r = OP$ designates the radius vector, θ corresponds to the "latitude," and η to the "longitude." The connection between the coördinates r, θ, η and the rectangular coördinates is given by the following relations:

$$x = ON = OM \cos \eta = OP \sin \theta \cos \eta = r \sin \theta \cos \eta \qquad (19a)$$

$$y = MN = r \sin \theta \sin \eta \qquad (19b)$$

$$z = PM = r \cos \theta, \qquad (19c)$$

where PM is the perpendicular from the point P to the plane OXY (the equatorial plane).

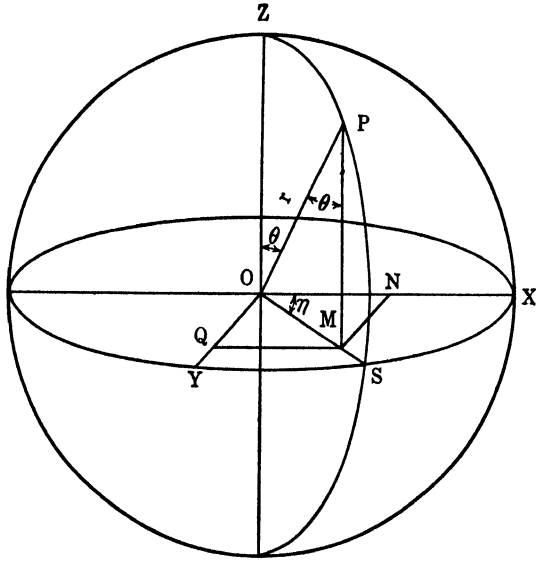


FIG. 22*a*. Illustrating relation between rectangular and spherical coördinates.

For the motion of an electron in the field due to two nuclei separated by a fixed distance, confocal elliptic coördinates are used, and in some problems still other systems of coördinates may be convenient.

Corresponding to each of the coördinate variables used to specify the position of a particle there will be a velocity component. With rectangular coördinates these components are $\dot{x}$, $\dot{y}$, and $\dot{z}$, and the velocity in the direction OP (Fig. 20) is determined from the relation

$$v^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2. \qquad (20)$$

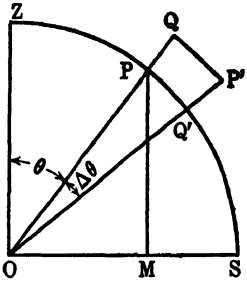
For the case of spherical coördinates it follows from equations (19) that

$$\dot{x} = r\cos\theta \cdot \dot{\theta} \cdot \cos\eta - r\sin\theta \cdot \sin\eta \cdot \dot{\eta} + \dot{r}\sin\theta \cdot \cos\eta$$
$$\dot{y} = r\cos\theta \cdot \dot{\theta} \cdot \sin\eta + r\sin\theta \cdot \cos\eta \cdot \dot{\eta} + \dot{r}\sin\theta \cdot \sin\eta$$
$$\dot{z} = -r\sin\theta \cdot \dot{\theta} + \dot{r}\cos\theta.$$

Hence

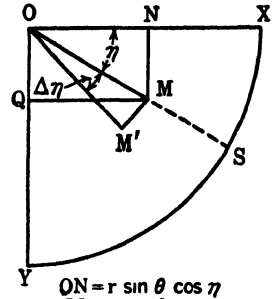
$$v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = \dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta \cdot \dot{\eta}^2. \tag{21}$$

This result may also be deduced readily from an inspection of Figs. 22*b* and 22*c*.[5]



P'Q=PQ'=rΔθ
PQ=P'Q'=Δr
OM=r sinθ

FIG. 22*b*. Projection on meridian plane.



ON=r sinθ cosη
OQ=r sinθ sinη
MM'=OM.Δη=r sinθΔη

FIG. 22*c*. Projection on equatorial plane.

Cylindrical and spherical coördinates are illustrations of a general class known as *orthogonal curvilinear* systems of coördinates which have certain common properties. Let these generalized coördinates be designated by q_1, q_2, and q_3. Then it will be found that the three sets of coördinate surfaces q_1 = constant, q_2 = constant, and q_3 = constant intersect at right angles. For each of these systems, the element of distance ds is given by a relation of the form

$$(ds)^2 = a_1(dq_1)^2 + a_2(dq_2)^2 + a_3(dq_3)^2, \tag{22}$$

[5] Figure 22*b* shows a projection of the meridian plane containing the right-angled triangle OMP, while Fig. 22*c* shows a similar projection of the equatorial plane containing the right-angled triangle ONM. It is evident that $r\Delta\theta \cdot \Delta r$ is the area of the element of $PQP'Q'$. From this it follows that $r\Delta\theta \cdot \Delta r \cdot r\sin\theta \cdot \Delta\eta$ is the element of volume. In rectangular coördinates, the element of volume is $\Delta x \cdot \Delta y \cdot \Delta z$. Hence, in the limit,

$$dxdydz = r^2\sin\theta drd\theta d\eta,$$

and it is seen that $r^2\sin\theta$ is the coefficient by which $drd\theta d\eta$ must be multiplied to convert it into an element of volume. See further remarks in Appendix III.

where a_1, a_2, and a_3 are coefficients involving one or more of the three generalized coördinates. Also the element of volume in terms of any orthogonal system of coördinate variables is given by

$$d\tau = dxdydz = \sqrt{a_1a_2a_3}\, dq_1dq_2dq_3, \tag{23}$$

and the coefficient $\sqrt{a_1a_2a_3}$ is known as the *discriminant* for the transformation from rectangular to orthogonal curvilinear coördinates.[6]

In general, the positions, at any instant of time, of all the particles constituting a given system may be specified by f variables, which we shall designate by the letters $q_1, q_2 \ldots q_f$. For a system of n particles the maximum value of f would be $3n$, but usually relations of constraints will exist between the particles, so that f will be less than $3n$. If these variables are so chosen that f is a minimum they are designated as *generalized coördinates* and f then corresponds to the *number of degrees of freedom of the system.* It is customary to designate these generalized coördinates by the letters $q_1, q_2 \ldots q_f$, and the corresponding *generalized velocities* are therefore $\dot{q}_1, \dot{q}_2 \ldots \dot{q}_f$.

In a system consisting of n particles, the rectangular coördinates of the ith particle will be expressed in terms of the generalized variables by functional relations of the form[7]

$$\left.\begin{aligned} x_i &= x_i(q_1, q_2 \ldots q_f) \\ y_i &= y_i(q_1, q_2 \ldots q_f) \\ z_i &= z_i(q_1, q_2 \ldots q_f) \end{aligned}\right\}. \tag{24}$$

Hence any arbitrary small variation in x_i will be given by a relation of the form

$$\delta x_i = \frac{\partial x_i}{\partial q_1}\delta q_1 + \frac{\partial x_i}{\partial q_2}\delta q_2 + \ldots + \frac{\partial x_i}{\partial q_f}\delta q_f, \tag{25}$$

with similar expressions for δy_i and δz_i. The partial differential coefficient $\partial x_i/\partial q_1$ is used in this relation instead of the more exact expression $\partial x_i(q_1, q_2 \ldots q_f)/\partial q_1$, and in equation (25) it will be observed that while δx_i designates a variation in the coördinate variable x_i, the symbol x_i in the partial differential coefficient represents a function of the f generalized coördinates $q_1 \ldots q_f$.

[6] For the derivation of equation (23) and discussion of these systems of coördinates, see Appendix III, and also the following references: Slater and Frank, "Introduction to Theoretical Physics," p. 200; Pauling and Wilson, "Introduction to Quantum Mechanics," p. 103 and Appendix IV; A. G. Webster, "The Dynamics of Particles and of Rigid, Elastic and Fluid Bodies," Chapter VIII.

[7] The following discussion is based on the excellent chapter, "Advanced Dynamics," in L. Page's "Introduction to Theoretical Physics."

It follows that the velocity components are given by relations of the form

$$\dot{x}_i = \frac{\partial x_i}{\partial q_1}\dot{q}_1 + \frac{\partial x_i}{\partial q_2}\dot{q}_2 + \ldots + \frac{\partial x_i}{\partial q_f}\dot{q}_f, \tag{26}$$

with similar expressions for $\dot{y}_1$ and $\dot{z}_i$.

Equation (26) shows that the expression for $\dot{x}_i$ ($\dot{y}_i$ or $\dot{z}_i$) is a *homogeneous linear function* of the generalized velocities, with coefficients which are functions of the q's. Since

$$T = \sum_i \left\{ \tfrac{1}{2}\mu_i(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) \right\}, \tag{27}$$

it follows from (26) that T is a *homogeneous quadratic function* of the q's.

From this fact may be deduced another important result by application of Euler's theorem for homogeneous functions. This theorem states that if u is a homogeneous function of the nth degree of any number (f) of variables $x_1, x_2 \ldots x_f$, then

$$nu = x_1 \cdot \frac{\partial u}{\partial x_1} + x_2 \cdot \frac{\partial u}{\partial x_2} + \ldots + x_f \cdot \frac{\partial u}{\partial x_f}.$$

Since T is a homogeneous function of the second degree in the generalized velocities, it follows from Euler's theorem that

$$2T = \dot{q}_1 \cdot \frac{\partial T}{\partial \dot{q}_1} + \dot{q}_2 \cdot \frac{\partial T}{\partial \dot{q}_2} + \ldots + \dot{q}_f \cdot \frac{\partial T}{\partial \dot{q}_f}. \tag{28}$$

4.4 Lagrange's Equations. In section 2 it was shown that Newton's second law of motion may be expressed in terms of rectangular coördinate variables by the set of equations (8). Can the laws of dynamics be obtained in forms which are independent of the particular nature of the coördinate system used? This was the question which Lagrange, the great French mathematician, answered in his "Mécanique analytique." As Lindsay and Margenau describe his achievement:[8]

> Lagrange proceeded from very general considerations, endeavoring to reduce mechanics to pure analysis and emancipate it from the connection with geometry which had been one of its outstanding characteristics as developed by Newton. With triumphant éclat Lagrange announced in the preface to his book that "there are no figures in this work," implying that all had been reduced to algebraic analysis (in the large sense).

We shall now consider the method used by Lagrange in deriving the laws of dynamics in terms of generalized coördinates.

[8] Lindsay and Margenau, "Foundations of Physics," p. 136.

In equation (9) let us replace δx_i by the expression derived in equation (25), and similarly for δy_i and δz_i, and for all the n particles. The resultant expression of d'Alembert's Principle is of the form

$$\sum_{i=1}^{n}\left[(\mu_i\ddot{x}_i - X_i)\left(\frac{\partial x_i}{\partial q_1}\cdot\delta q_i + \ldots + \frac{\partial x_i}{\partial q_f}\cdot\delta q_f\right)\right] = 0. \tag{29}$$

Evidently,

$$\ddot{x}_i\frac{\partial x_i}{\partial q_k} = \frac{d}{dt}\left(\dot{x}_i\cdot\frac{\partial x_i}{\partial q_k}\right) - \dot{x}_i\cdot\frac{d}{dt}\left(\frac{\partial x_i}{\partial q_k}\right), \tag{30}$$

and from equation (26) it is seen that

$$\frac{\partial x_i}{\partial q_k} = \frac{\partial \dot{x}_i}{\partial \dot{q}_k}.$$

Hence,

$$\frac{d}{dt}\left(\dot{x}_i\cdot\frac{\partial x_i}{\partial q_k}\right) = \frac{d}{dt}\left(\dot{x}_i\cdot\frac{\partial \dot{x}_i}{\partial \dot{q}_k}\right) = \frac{1}{2}\cdot\frac{d}{dt}\left\{\frac{\partial(\dot{x}_i^2)}{\partial \dot{q}_k}\right\}. \tag{31}$$

Also

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial x_i}{\partial q_k}\right) &= \frac{\partial^2 x_i}{\partial q_1 \partial q_k}\cdot\dot{q}_1 + \frac{\partial^2 x_i}{\partial q_2 \partial q_k}\cdot\dot{q}_2 + \ldots \\ &= \frac{\partial}{\partial q_k}\left(\frac{\partial x_i}{\partial q_1}\cdot\dot{q}_1 + \frac{\partial x_i}{\partial q_2}\cdot\dot{q}_2 + \ldots\right) \\ &= \frac{\partial \dot{x}_i}{\partial q_k} \text{ from equation (26).}\end{aligned}$$

Therefore,

$$\dot{x}_i\frac{d}{dt}\left(\frac{\partial x_i}{\partial q_k}\right) = \dot{x}_i\cdot\frac{\partial \dot{x}_i}{\partial q_k} = \frac{\partial}{\partial q_k}\left(\frac{1}{2}\dot{x}_i^2\right). \tag{32}$$

Substituting from equations (31) and (32) in (30) it follows that

$$\ddot{x}_i\cdot\frac{\partial x_i}{\partial q_k} = \frac{d}{dt}\left(\frac{1}{2}\cdot\frac{\partial \dot{x}_i^2}{\partial \dot{q}_k}\right) - \frac{\partial}{\partial q_k}\left(\frac{1}{2}\dot{x}_i^2\right).$$

Consequently,

$$\mu_i\ddot{x}_i\delta x_i = \sum_f\left\{\frac{d}{dt}\cdot\frac{\partial}{\partial \dot{q}_k}\left(\frac{1}{2}\mu_i\dot{x}_i^2\right) - \frac{\partial}{\partial q_k}\left(\frac{1}{2}\mu_i\dot{x}_i^2\right)\right\}\delta q_k. \tag{33}$$

Now $(\frac{1}{2})\mu_i\dot{x}_i^2$ is the kinetic energy along the x-axis of the ith particle. If we write down relations similar to equation (33) for $\mu_i\ddot{y}_i\delta y_i$ and $\mu_i\ddot{z}_i\delta z_i$ and sum up the right-hand sides of these equations, it is evident

that the summation of the terms similar to the first term on the right-hand side of (33) corresponds to

$$\sum \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right),$$

while the summation of the terms similar to the second term corresponds to $\sum \partial T/\partial q_k$. Hence, we derive the result,

$$\sum_i [(\mu_i \ddot{x}_i - X_i)\delta x_i + (\mu_i \ddot{y}_i - Y_i)\delta y_i + (\mu_i \ddot{z}_i - Z_i)\delta z_i]$$

$$= \sum_f \left\{ \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} + \frac{\partial V}{\partial q_k} \right\} \delta q_k = 0, \tag{34}$$

since

$$X_i \delta x_i + Y_i \delta y_i + Z_i \delta z_i = \delta V = \sum \frac{\partial V}{\partial q_k} \cdot \delta q_k.$$

Since V is *not* a function of the $\dot{q}$'s, it does not alter the sense of equation (34) if we write it in the form

$$\sum_f \left\{ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} \right\} \delta q_k = 0, \tag{35}$$

where $L = T - V$.

Now δq_1, δq_2 . . . δq_f are arbitrary variations. Consequently, equation (35) cannot be valid unless each of the f expressions in the brackets on the left-hand side vanishes identically. That is, *for each of the* f *generalized coördinates* there is valid an equation of the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0. \tag{36}$$

This is known as *Lagrange's equation*, and the function $L = L(q_k, \dot{q}_k)$ is known as the *Lagrangian* or *kinetic potential function.*

From the method of derivation, it is evident that the Lagrangian equations are valid for any coördinate system, as long as the number of coördinates corresponds to the total number of degrees of freedom of the system of particles.

For the simplest case, that of the motion of a single particle along the x-axis, equation (36) becomes

$$\frac{d}{dt}\left(\frac{dT}{d\dot{x}}\right) - \left(\frac{dT}{dx} + X\right) = 0,$$

where $X = -dV/dx$.

Since $T = \mu\dot{x}^2/2$, $dT/d\dot{x} = \mu\dot{x}$ = momentum, and $dT/dx = 0$, Lagrange's equation assumes the simple form

$$\mu\ddot{x} - X = 0,$$

which is Newton's second law of motion.

4.5 Motion of Electron in Bohr Atom in Absence of Quantizing Conditions.[9] By means of the Lagrange equations it is possible to derive the equations of motion for an electron in the field due to a nucleus of charge $+Ne$. This is a typical case of central field motion and is described most conveniently in terms of the spherical coördinates r, θ, and η. The equations for the transformation from rectangular coördinates have been given in (19), and the corresponding expression for the kinetic energy of the electron is given, in accordance with equation (21), by the relation

$$T = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta \cdot \dot{\eta}^2). \tag{37}$$

The potential energy function, as shown in section 2, is given by the relation

$$V = -\frac{Ne^2}{r}. \tag{15}$$

Hence,

$$\frac{\partial L}{\partial \dot{r}} = \frac{\partial T}{\partial \dot{r}} = \mu\dot{r} \tag{38}$$

$$\frac{\partial L}{\partial \dot{\theta}} = \frac{\partial T}{\partial \dot{\theta}} = \mu r^2\dot{\theta} \tag{39}$$

$$\frac{\partial L}{\partial \dot{\eta}} = \mu r^2 \sin^2\theta \cdot \dot{\eta} \tag{40}$$

$$\frac{\partial L}{\partial r} = \frac{\partial T}{\partial r} - \frac{\partial V}{\partial r} = \mu r(\dot{\theta}^2 + \sin^2\theta \cdot \dot{\eta}^2) - \frac{Ne^2}{r^2} \tag{41}$$

$$\frac{\partial L}{\partial \theta} = \mu r^2 \sin\theta \cos\theta \cdot \dot{\eta}^2 \tag{42}$$

$$\frac{\partial L}{\partial \eta} = 0 \tag{43}$$

[9] The discussion in this section follows that given by Lindsay and Margenau, "Foundations of Physics," pp. 139–140.

Consequently, the Lagrange equations are as follows:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = \mu\ddot{r} + \frac{Ne^2}{r^2} - \mu r(\dot{\theta}^2 + \sin^2\theta \cdot \dot{\eta}^2) = 0, \tag{44}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt}(\mu r^2\dot{\theta}) - \mu r^2 \sin\theta\cos\theta \cdot \dot{\eta}^2 = 0, \tag{45}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\eta}}\right) - \frac{\partial L}{\partial \eta} = \frac{d}{dt}(\mu r^2 \sin^2\theta \cdot \dot{\eta}) = 0. \tag{46}$$

In consequence of the last equation,

$$\mu r^2 \sin^2\theta \cdot \dot{\eta} = C, \tag{47}$$

where C is a constant of integration.

Eliminating $\dot{\eta}$ from equations (44) and (45) by means of (47) we obtain the two *equations of motion*

$$\mu\ddot{r} + \frac{Ne^2}{r^2} - \mu r\left(\dot{\theta}^2 + \frac{C^2}{\mu^2 r^4 \sin^2\theta}\right) = 0, \tag{48}$$

and

$$\frac{d}{dt}(\mu r^2\dot{\theta}) - \frac{C^2\cos\theta}{\mu r^2 \sin^3\theta} = 0. \tag{49}$$

The fact that the problem is thus reduced to one involving only two degrees of freedom shows that the motion occurs in a *plane* and may be represented in terms of the variables r and θ. Consequently, we can take the plane as that for which $\eta = 0$, and we have $\dot{\eta} = 0$ and $C = 0$, so that equations (48) and (49) assume the simpler forms

$$\mu\ddot{r} + \frac{Ne^2}{r^2} - \mu r\dot{\theta}^2 = 0, \tag{50}$$

and

$$\frac{d}{dt}(\mu r^2\dot{\theta}) = 0. \tag{51}$$

The last equation leads to the very important conclusion

$$\mu r^2\dot{\theta} = \mu r^2\frac{d\theta}{dt} = \alpha, \tag{52}$$

where α is a constant, and equation (52) states that the area described per unit time by the radius vector, which is equal to $(\frac{1}{2})r^2\dot{\theta}$, is a constant. This constitutes one of Kepler's laws for planetary orbits.

Substituting this result in equation (50) we obtain *the second-order differential equation for the radial motion of the electron* in the form

$$\frac{d^2r}{dt^2} + \frac{Ne^2}{\mu r^2} - \frac{\alpha^2}{\mu^2 r^3} = 0. \tag{53}$$

Multiplying through by $2dr/dt$, we obtain the equation

$$\frac{d}{dt}\left(\frac{dr}{dt}\right)^2 - \frac{2Ne^2}{\mu}\frac{d}{dt}\left(\frac{1}{r}\right) + \frac{\alpha^2}{\mu^2}\cdot\frac{d}{dt}\left(\frac{1}{r^2}\right) = 0,$$

which, on integration, gives the equation

$$\left(\frac{dr}{dt}\right)^2 - \frac{2Ne^2}{\mu r} + \frac{\alpha^2}{\mu^2 r^2} = \beta,$$

where β is an integration constant.

Hence,

$$\frac{dr}{dt} = \sqrt{\beta + \frac{2Ne^2}{\mu r} - \frac{\alpha^2}{\mu^2 r^2}}. \tag{54}$$

Combining this with equation (52), and thus eliminating dt, we obtain the *differential equation for the orbit* in the form

$$\frac{d\theta}{dr} = \frac{\alpha}{\mu r^2\sqrt{\beta + \frac{2Ne^2}{\mu r} - \frac{\alpha^2}{\mu^2 r^2}}}. \tag{55}$$

Introducing the variable

$$z = \frac{\alpha}{\mu r} - \frac{Ne^2}{\alpha},$$

it is readily shown that the integral of equation (55) is

$$\left(\frac{\alpha^2}{Ne^2\mu}\right)\frac{1}{r} = 1 + \sqrt{1 + \frac{\alpha^2\beta}{N^2e^4}}\cdot\cos(\theta + \gamma), \tag{56}$$

where γ is an integration constant which may be equated to zero, since it involves only the location of the polar axis with respect to the axis of reference.

Comparing equation (56) with the polar equation[10] for a conic section in the form

$$\frac{l}{r} = 1 + \epsilon\cos\theta, \tag{57}$$

[10] See supplementary note 1 for derivation of equation of conic section in terms of polar coördinates.

where l is the semi-parameter, and ϵ the eccentricity, it is seen that equation (56) represents a similar curve for which

$$l = \frac{\alpha^2}{Ne^2\mu} \tag{58}$$

and

$$\epsilon = \sqrt{1 + \frac{\alpha^2\beta}{N^2e^4}}. \tag{59}$$

For $\epsilon < 1$, the curve represents an ellipse; for $\epsilon = 0$, a circle; for $\epsilon > 1$, the curve represents a hyperbola; while for $\epsilon = 1$, it corresponds to a parabola. It is therefore necessary to determine the physical significance of the integration constant β.

Now

$$E = T + V = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{Ne^2}{r}.$$

Substituting from equations (52) and (54) it follows that

$$E = \frac{\mu}{2}\left(\beta + \frac{2Ne^2}{\mu r} - \frac{\alpha^2}{\mu^2r^2} + \frac{\alpha^2}{\mu^2r^2}\right) - \frac{Ne^2}{r} = \frac{\mu\beta}{2}.$$

That is,

$$\beta = \frac{2E}{\mu}. \tag{60}$$

Since E is negative for a central orbit,[11] it follows that β is negative, and therefore ϵ, as defined by equation (59), is less than or equal to 1. That is, the orbit is an ellipse or circle.

Let $W = -E$, so that W is a positive magnitude. Then it follows from equations (58), (59), and (60) that

$$1 - \epsilon^2 = \frac{2\alpha^2W}{N^2e^4\mu} = \frac{2Wl}{Ne^2}.$$

But $2l/(1 - \epsilon^2) = 2a$, the major axis of the ellipse.

Therefore

$$E = -W = -\frac{Ne^2}{2a}. \tag{61}$$

That is, the *energy depends only upon the magnitude of the major axis* of the ellipse and is independent of the eccentricity.

[11] See supplementary note 2 for discussion of this statement.

In section 8 it will be shown how Bohr introduced a quantizing condition and thus derived the conclusion that the value of $2a$, the major axis, cannot vary continuously (corresponding to a continuous variation in E) but that the electron can revolve only in a series of discrete orbits, corresponding to a series of discrete values of E.

From equation (52) the period of revolution (τ) of the electron in its orbit may be deduced as follows:

The area of an ellipse, for which $2a$ is the major and $2b$ the minor axis, is given by πab. It follows from equation (52) that

$$\frac{\alpha}{2\mu} = \frac{\pi ab}{\tau};$$

that is,

$$\tau = \frac{2\pi ab \cdot \mu}{\alpha}. \tag{62}$$

But

$$b = a\sqrt{1 - \epsilon^2},$$

and from equations (59) and (60) it follows that

$$\frac{1 - \epsilon^2}{\alpha^2} = -\frac{2E}{\mu N^2 e^4} = \frac{2W}{\mu}\left(\frac{1}{Ne^2}\right)^2. \tag{63}$$

Substituting for b and α in equation (62) it follows that

$$\tau = 2\pi a^2 \mu \sqrt{\frac{2W}{\mu}} \cdot \frac{1}{Ne^2}.$$

Substituting for W from equation (61), we obtain the two relations

$$\tau = 2\pi a^{\frac{3}{2}} \sqrt{\frac{\mu}{Ne^2}} \tag{64a}$$

and

$$\tau = \frac{2\pi Ne^2 \sqrt{\mu}}{2\sqrt{2} \cdot W^{\frac{3}{2}}}. \tag{64b}$$

Equation (64a) corresponds to Kepler's third law for planetary orbits which states that the square of the period of revolution varies as the cube of the semi-major axis. Equation (64b) gives τ in terms of the total energy $E = -W$.

In deducing these relations the assumption has been introduced implicitly that the nucleus is at rest. Actually both the nucleus and

electron move in elliptic orbits about their common center of gravity. The motion of the electron is then described in terms of a particle about a fixed center if the " reduced " mass μ_0, defined by the relation

$$\frac{1}{\mu_0} = \frac{1}{\mu} + \frac{1}{M},$$

is used, and the radius vector r is defined by the relations

$$r_1 = r\left(\frac{M}{\mu + M}\right); \; r_2 = r\left(\frac{\mu}{\mu + M}\right),$$

where $r = r_1 + r_2$. In these relations μ and M designate the mass of electron and nucleus respectively, while r_1 and r_2 represent the distances of electron and nucleus respectively from their common center of gravity.[12]

It is evident that for M very large compared to μ the distance r_2 is vanishingly small, so that $r \approx r_1$, and $\mu_0 \approx \mu$.[13]

4.6 Canonical Equations. Since each Lagrangian equation is a differential equation of the *second order*, it is more convenient in dynamical problems to replace these f equations (corresponding to f degrees of freedom) by $2f$ differential equations of the *first order*.

In order to derive these equations we introduce a new and very important variable p_k, the *generalized momentum*, which is defined by the differential equation

$$p_k = \frac{\partial L}{\partial \dot{q}_k} = \frac{\partial T}{\partial \dot{q}_k}. \tag{65}$$

Since V is not a function of $\dot{q}_k$ but $L = T - V$ is a function of $q_1, q_2 \ldots q_f, \dot{q}_1, \dot{q}_2, \ldots \dot{q}_f$, and in the most general case of t as well, we indicate this by the relation

$$L = L(q_k, \dot{q}_k, t). \tag{66}$$

In terms of the new variable, Lagrange's equations assume the simplified form

$$\frac{d}{dt}(p_k) = \dot{p}_k = \frac{\partial L}{\partial q_k}. \tag{67}$$

[12] For detailed discussion of this case the student should consult Ruark and Urey, "Atoms, Molecules, and Quanta," and Pauling and Wilson, "Introduction to Quantum Mechanics."

[13] The sign $\approx$ is used to designate approximate, but not exact, equality.

From equation (66) it follows that we can write the total differential of L with respect to the three sets of variables in the form

$$dL = \sum_f \frac{\partial L}{\partial q_k} \cdot dq_k + \sum_f \frac{\partial L}{\partial \dot{q}_k} \cdot d\dot{q}_k + \frac{\partial L}{\partial t} \cdot dt.$$

Incorporating equations (65) and (67), the last equation becomes

$$dL = \sum_f \dot{p}_k dq_k + \sum_f p_k d\dot{q}_k + \frac{\partial L}{\partial t} \cdot dt. \tag{68}$$

Now let us introduce a function H, defined by the relation[14]

$$H = \sum_f p_k \dot{q}_k - L. \tag{69}$$

Then it follows from (68) and (69) that

$$\begin{aligned} dH &= \sum_f \left\{ p_k d\dot{q}_k + \dot{q}_k dp_k - \dot{p}_k dq_k - p_k d\dot{q}_k \right\} - \frac{\partial L}{\partial t} \cdot dt \\ &= \sum_f \left\{ \dot{q}_k dp_k - \dot{p}_k dq_k \right\} - \frac{\partial L}{\partial t} \cdot dt. \end{aligned} \tag{70}$$

Since dH is an exact differential, equation (70) indicates that H is a function of the independent variables p_k, q_k, and t, that is,

$$H = H(p_k, q_k, t).$$

Hence,

$$dH = \sum_f \frac{\partial H}{\partial p_k} \cdot dp_k + \sum_f \frac{\partial H}{\partial q_k} \cdot dq_k + \frac{\partial H}{\partial t} \cdot dt. \tag{71}$$

Comparing coefficients of dp_k, dq_k, and dt in equations (70) and (71), we deduce the three first-order *partial* differential equations

$$\frac{\partial H}{\partial p_k} = \dot{q}_k = \frac{dq_k}{dt}, \tag{72}$$

$$\frac{\partial H}{\partial q_k} = -\dot{p}_k = -\frac{dp_k}{dt}, \tag{73}$$

and

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{74}$$

[14] This is known as a *Legendre transformation.*

These equations express the equations of motion in the so-called *canonical* form, and p_k and q_k are designated *canonically conjugated* variables. The function $H = H(p_k, q_k)$, in which t does not enter as an independent variable, is known as the *Hamiltonian*, and we shall now proceed to determine the physical significance of this function.

In this case, $\partial H/\partial t = 0$, and we derive from equation (71) the relation

$$\frac{dH}{dt} = \sum_f \frac{\partial H}{\partial p_k} \cdot \dot{p}_k + \sum_f \frac{\partial H}{\partial q_k} \cdot \dot{q}_k,$$

which in virtue of the canonical equations (72) and (73) becomes

$$\frac{dH}{dt} = \sum_f \dot{q}_k \dot{p}_k - \sum_f \dot{p}_k \dot{q}_k = 0.$$

Hence, $H(p_k, q_k) = constant$, is a *first integral* of the differential equations of motion (72) and (73).

Furthermore, by the definition of H, as given in (69),

$$H = \sum_f p_k \dot{q}_k - T + V. \tag{75}$$

But, as deduced in equation (28),

$$2T = \sum_f \dot{q}_k \cdot \frac{\partial T}{\partial \dot{q}_k} = \sum_f \dot{q}_k p_k. \tag{76}$$

Hence, equation (75) becomes

$$H = 2T - T + V = T + V = E. \tag{77}$$

Thus, *for a conservative system* $H = H(p_k, q_k)$ *is equal to the tota energy.*

It is evident, from the method used in the derivation of the canonica equations, that if we have two sets of canonically conjugated variables q_k, p_k and Q_k, P_k, for which $H(p_k, q_k) = H(Q_k, P_k) = E$, then equa tions (72) and (73) are equally valid for both sets of variables. Thi conclusion leads to a theory of *contact transformations* which is of grea importance in both classical and quantum mechanics.

4.7 Hamilton's Principle and the Principle of Least Action. Fror Lagrange's equations it is possible to deduce two principles which ar of extremely great importance in classical dynamics and have thei counterpart in quantum mechanics.

By the methods of the calculus of variation it is shown that th Lagrange equations (36) are a necessary condition which must b

satisfied in order that the integral

$$S = \int_{t_0}^{t_1} L dt \tag{78}$$

shall be a maximum or minimum (that is, a *stationary value* or *extremum*). For the integral S it can be shown that it is actually a minimum. This conclusion is usually described as *Hamilton's Principle* and is stated as follows.

If we compare a dynamical path (that is, one which proceeds in accordance with the laws of dynamics) with *varied paths which have the same termini (in the configuration space) and are described in the same time*, then the time integral of the Lagrangian function has a stationary value for the dynamical path, so that

$$\delta \int_{t_0}^{t_1} L dt = \delta \int_{t_0}^{t_1} (T - V) dt = 0, \tag{79}$$

where δ denotes that the path is to be varied in any arbitrary manner, subject, however, to the conditions mentioned above.

Since the integral S, defined in equation (78), has a definite value which depends only on the initial and final values of the coördinates (q_1^0, q_2^0 . . . q_f^0 and q_1^1, q_2^1 . . . q_f^1) and on the time interval $t_1 - t_0$, it follows that

$$\int_{t_0}^{t_1} L dt = S(q_k^1, q_k^0, t_1, t_0), \tag{80}$$

where S is known as *Hamilton's Function.*

For conservative systems, the energy is a constant. Now let us compare two infinitely near paths, for *each of which the energy is the same, and which are described in the same time interval* $t_1 - t_0$. Since each path is a natural one, and we shall assume that both have the same initial coördinates, it is evident that the final coördinates cannot be the same for both.

From Hamilton's Principle it follows that for a natural path

$$\delta \int_{t_0}^{t_1} L dt + \delta \int_{t_0}^{t_1} E dt = 0.$$

But $L = T - V$, and $E = T + V$. Hence this relation may be expressed in the form

$$\delta A = \delta \int_{t_0}^{t_1} 2T dt = 0, \tag{81}$$

where A is a function, known as the *Characteristic Function* or *Action*,

which must be an extremum (actually a minimum) for conservative systems. The relation stated in the last equation is known as the *Principle of Least Action.*

This relation can also be stated in another form, which proved convenient in the classical theory of atomic mechanics.

Since

$$T = \sum_i \tfrac{1}{2}\mu_i v_i^2,$$

where v_i is the velocity of the ith particle of mass u_i, it follows that

$$A = \sum \int_{t_0}^{t_1} \mu_i v_i^2 dt = \sum_i \int_{t_0}^{t_1} \mu_i v_i \cdot v_i dt = \sum_i \int_{s_0}^{s_1} \mu_i v_i ds_i, \tag{82}$$

where ds_i denotes an element of the path. This relation shows that the Action corresponds to the sum of the line integrals of the momenta taken over the total path from the initial point s_0 to the final point s_1.

For a single particle

$$\tfrac{1}{2}\mu v^2 = T = E - V,$$

and therefore

$$\mu v = \sqrt{2\mu(E - V)}.$$

Hence,

$$A = \int_{s_0}^{s_1} \sqrt{2\mu(E - V)} \cdot ds. \tag{83}$$

If V is a function of f-coördinates, then ds is an element of " path " in the f-dimensional space.

4.8 The Wilson-Sommerfeld Quantum Conditions. Equation (82) can also be deduced by making use of equation (76). Combining (76) with (81) it follows that

$$A = \int_{t_0}^{t_1} 2T dt = \sum_f \int_{t_0}^{t_1} p_k \dot{q}_k dt = \sum_f \int_{q_k^0}^{q_k^1} p_k dq_k. \tag{84}$$

Since the value of A depends only upon the initial and final values of the coördinates and the energy E, it is an *integral invariant.* For any system for which E is given, the value of A taken between any two points in the f-dimensional space (configuration space) must be a constant and therefore *independent of the particular type of coördinate system.*

In particular, if we consider a periodic type of orbit, the value of the Action, or line integral of momenta, taken over a complete period of revolution, must be a constant.

Wilson and Sommerfeld therefore postulated that for atomic systems this integral, known as the *Action integral*, and designated by J_k (one for each degree of freedom or coördinate variable) *must be equal to an integral of Planck's constant h* (which has the dimensions of action). That is, the Wilson-Sommerfeld *quantum conditions* are of the form

$$J_k = \oint p_k dq_k = n_k h. \tag{85}$$

The circle around the sign of integration indicates integration over the orbit for a complete period τ, where $\nu = 1/\tau$ is the frequency of the particular motion, and n_k is an integer, which may have a different value for each coördinate variable.

Thus, for an elliptic orbit of the electron in the hydrogen-like atom, the quantum conditions are given by

$$J_\theta = \oint p_\theta d\theta = n_\theta h \tag{86}$$

$$J_r = \oint p_r dr = n_r h. \tag{87}$$

In this case, as discussed in section 5,

$$E = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{Ne^2}{r}$$

Therefore,

$$p_\theta = \frac{\partial T}{\partial \dot{\theta}} = \mu r^2 \dot{\theta},$$

and

$$p_r = \mu \dot{r}.$$

Hence,

$$E = H(p_\theta, p_r) = \frac{p_r^2}{2\mu} + \frac{p_\theta^2}{2\mu r^2} - \frac{Ne^2}{r}, \tag{88}$$

and applying the canonical equations corresponding to θ,

$$\frac{\partial H}{\partial p_\theta} = \dot{\theta} \tag{i}$$

$$\frac{\partial H}{\partial \theta} = 0 = -\dot{p}_\theta, \tag{ii}$$

that is,

$$p_\theta = \text{constant} = \alpha,$$

as deduced in equation (52).

Hence,

$$J_\theta = \oint p_\theta d\theta = 2\pi\alpha = n_\theta h.$$

In the case of the coördinate r,

$$\frac{\partial H}{\partial p_r} = \dot{r} = \frac{p_r}{\mu} \tag{iii}$$

$$\frac{\partial H}{\partial r} = -\dot{p}_r = -\frac{\alpha^2}{\mu r^3} + \frac{Ne^2}{r^2}. \tag{iv}$$

The last equation is identical with equation (53), and the solution has been given in section 5. In order to apply the quantum condition it is necessary to evaluate the integral

$$J_r = n_r h = 2\int_{r_0}^{r_m} p_r dr, \tag{v}$$

where r_0 and r_m denote the minimum and maximum values respectively of the radius vector.

Now from equations (54) and (60) we have the relation

$$\begin{aligned} p_r = \mu r &= \sqrt{2\mu E + 2Ne^2\frac{\mu}{r} - \frac{\alpha^2}{r^2}} \\ &= \sqrt{2\mu E + 2Ne^2\frac{\mu}{r} - \frac{J_\theta^2}{4\pi^2 r^2}} \\ &= \sqrt{Q}. \end{aligned}$$

Let us consider the significance of the function Q. It is evident that

$$\begin{aligned} \frac{r^2 Q}{2\mu W} &= -r^2 + \frac{Ne^2}{W}\cdot r - \frac{\alpha^2}{2\mu W} \\ &= -r^2 + 2ar - \frac{\alpha^2}{2\mu}\cdot\frac{2a}{Ne^2} \\ &= -r^2 + \{a(1+\epsilon) + a(1-\epsilon)\}\, r - a^2(1-\epsilon^2). \end{aligned}$$

In deducing these relations use has been made of equations (61) and (63). Now from the properties of the ellipse it follows that

$$a(1+\epsilon) = r_m;\ a(1-\epsilon) = r_0.$$

Consequently,

$$Q = \frac{2\mu W}{r^2}(r - r_0)(r_m - r), \tag{89}$$

and it is seen that $Q = 0$ for $r = r_0$ and $r = r_m$.

The evaluation of

$$J_r = 2\int_{r_0}^{r_1} Q \cdot dr$$

is obtained most readily by application of the theory of functions of complex variables,[15] and the result is

$$J_r = 2\pi i\left(\frac{iJ_\theta}{2\pi} + \frac{Ne^2\mu}{\sqrt{-2\mu E}}\right).$$

Hence

$$E_n = -\frac{2\pi^2N^2e^4\mu}{(J_r + J_\theta)^2} = -\frac{2\pi^2N^2e^4\mu}{(n_r + n_\theta)^2h^2} \tag{90}$$

$$= -\frac{RchN^2}{n^2},$$

where R = Rydberg constant (in terms of wave numbers)

$$= \frac{2\pi^2e^4\mu}{ch^3},$$

E_n = energy of atomic system in state of total quantum number n,

$n = n_r + n_\theta$ = an integer,

and c = velocity of light.

More strictly, the problem of the hydrogen-like atom should have been treated as a system of three dimensions, that is, by use of spherical polar coördinates.[16]

When the three quantum conditions (involving an additional one for the angle η) are applied, it is found that the value of E is the same as that derived in (90) with the only difference that n, the total quantum number, is the sum of three quantum numbers n_r, n_θ, and n_η. That is, E is independent of the individual values of these three numbers and *depends only on the value of their sum.* Thus, for the system in terms of

[15] See A. Sommerfeld, "Atombau und Spektrallinien" (Ed. 1924), pp. 772–779, and J. H. Van Vleck, "Quantum Principles and Line Spectra," pp. 193–196. See also Peirce's tables, formulae 161, 183, and 187. This problem is also discussed by L. Page, " Introduction to Theoretical Physics," pp. 567–570, in terms of the Action and Angle variables.

[16] See J. H. Van Vleck, "Quantum Principles," p. 193 et seq., and M. Born's "Atommechanik," Chapter III, for a comprehensive discussion of this problem. In classical mechanics the solution is more conveniently obtained by application of the Hamilton-Jacobi partial differential equation.

two variables, we have the possibilities indicated in the following table, in which n_θ is replaced by the more customary symbol k.

n	n_r	k	*Type of orbit*	$\frac{b}{a}$
1	0	1	Circular	1
2	0	2	Circular	1
	1	1	Elliptic	0.5
3	0	3	Circular	1
	1	2	Elliptic	0.67
	2	1	Elliptic	0.33

and similarly for larger values of n.

Since, as shown in section 5, the total energy is given by the relation

$$E_n = -\frac{Ne^2}{2a_n},$$

it follows from equation (90) that the semi-major axis of the orbit in the state of quantum number n is given by

$$\begin{aligned} a_n &= -\frac{Ne^2}{2E_n} \\ &= \frac{n^2}{N} \cdot \frac{h^2}{4\pi^2 e^2 \mu} \\ &= \frac{n^2}{N} \cdot a_1, \end{aligned} \tag{91}$$

where a_1 is the radius of the orbit in the normal state of the hydrogen atom. This is usually designated by the symbol a_0, where

$$a_0 = \frac{h^2}{4\pi^2 e^2 \mu} = 0.5283 \times 10^{-8} \text{ cm}. \tag{92}$$

Making use of the relations deduced in section 5, it is evident that

$$\begin{aligned} J_\theta &= 2\pi\alpha = 2\pi\sqrt{Ne^2\mu l} \\ &= 2\pi\sqrt{Ne^2\mu a(1-\epsilon^2)}. \end{aligned}$$

Hence,

$$\frac{J_\theta^2}{(J_r + J_\theta)^2} = \frac{k^2}{n^2} = 1 - \epsilon^2, \tag{93a}$$

and

$$\frac{k}{n} = \frac{b}{a}, \tag{93b}$$

where b is the semi-minor axis of the ellipse.

The last relation follows from equation (vii) in supplementary note 1. The values of b/a given in the above table for the different orbits show that, with increase in the value of $n_r = n - k$, the orbits become more elliptical.

In the cases $n = 2, n = 3$, and so forth, it is thus possible to have two or more orbits with the same value of E. Such states are known as *degenerate*, and in the first case we speak of a twofold degeneracy, in the second, of threefold degeneracy, and so on. It is only in the presence of electrostatic or magnetic fields that this degeneracy is removed. We shall meet with a similar situation when discussing the hydrogen-like atom in wave mechanics.

4.9 Mean Values for Central Orbits. In some of the problems involving central orbits it is of importance to calculate the mean value of a given magnitude M. By definition the mean value is given by the relation

$$\overline{M} = \frac{1}{\tau}\int_0^\tau M dt, \tag{94}$$

where τ is the period of revolution, and a bar across the symbol designates the mean value.

For the purpose of comparison with the results deduced by application of the S. equation it is of importance, in connection with the problem of the hydrogen-like atom, to calculate mean values of r and $1/r$ for the electronic orbits in the Bohr atom.[17]

From equation (94) we have the relations

$$\bar{r} = \frac{1}{\tau}\int_0^\tau r dt \tag{95}$$

$$\left(\frac{1}{r}\right) = \frac{1}{\tau}\int_0^\tau \frac{dt}{r}. \tag{96}$$

Now if in equation (54) we introduce the values of β and α^2 given by equations (60) and (58) respectively, equation (95) assumes the form

$$\bar{r} = \frac{2\sqrt{\mu}}{\tau}\int_{r_0}^{r_m} \frac{r dr}{\sqrt{\frac{2Ne^2}{r} - \frac{Ne^2 l}{r^2} - 2W}}, \tag{97a}$$

where r_0 and r_m are the minimum and maximum values of r, and $W = -E$. Proceeding as in the deduction of equation (89), we obtain

[17] For more comprehensive discussion see M. Born, "Atommechanik," p. 163 et seq.

the relation

$$\bar{r} = \frac{2\sqrt{\mu}}{\tau}\int_{r_0}^{r_m} \frac{r^2 dr}{\sqrt{2W(r - r_0)(r_m - r)}}. \tag{97b}$$

Let us now introduce the relation for r in terms of the *eccentric anomaly* u (which is derived in supplementary note 1), that is,

$$r = a(1 - \epsilon \cos u).$$

Then $$r_m = a(1 + \epsilon),$$

and $$r_0 = a(1 - \epsilon).$$

Therefore equation (97*b*) becomes

$$\bar{r} = \frac{2}{\tau}\sqrt{\frac{\mu}{2W}}\int_0^{\pi} \frac{a^2(1 - \epsilon \cos u)^2 \cdot a\epsilon \sin u \cdot du}{a\epsilon \sin u}$$

$$= \frac{2}{\tau}\sqrt{\frac{\mu}{2W}}\int_0^{\pi} a^2(1 - \epsilon \cos u)^2 du$$

$$= \frac{2a^2}{\tau}\sqrt{\frac{\mu}{2W}}\left(\pi + \frac{\epsilon^2\pi}{2}\right)$$

$$= 2a^{\frac{5}{2}}\frac{\pi}{\tau}\sqrt{\frac{\mu}{Ne^2}}\left(1 + \frac{\epsilon^2}{2}\right).$$

Introducing the value of τ deduced in equation (64*a*), it follows that

$$\bar{r} = a\left(1 + \frac{\epsilon^2}{2}\right),$$

and substituting for a from equation (91), and for ϵ^2 from equation (93*a*), it follows that

$$\bar{r} = \frac{n^2 a_0}{N}\left[1 + \frac{1}{2}\left(1 - \frac{k^2}{n^2}\right)\right]. \tag{98}$$

We shall now consider equation (96) which defines the mean value of $1/r$. Proceeding as in the previous case, we obtain the relation

$$\frac{1}{\tau}\int_0^{\tau} \frac{dt}{r} = \frac{2}{\tau}\sqrt{\frac{\mu}{2W}}\int_0^{\pi} \frac{dr}{\sqrt{(r - r_0)(r_m - r)}},$$

$$= \frac{2}{\tau}\sqrt{\frac{\mu}{2W}}\int_0^{\pi} \frac{a\epsilon \sin u \cdot du}{a\epsilon \sin u}$$

$$= \frac{2\pi}{\tau}\sqrt{\frac{\mu}{2W}} = \frac{1}{a^{\frac{3}{2}}}\sqrt{\frac{Ne^2}{2W}},$$

that is,

$$\left(\frac{1}{r}\right) = \frac{1}{a}. \tag{99}$$

This conclusion is obvious from the fact that the mean potential energy is given by

$$\bar{V} = -\frac{Ne^2}{2a},$$

and is therefore independent of b.

SUPPLEMENTARY NOTE 1

EQUATION OF ELLIPSE IN POLAR COÖRDINATES

In Fig. 23, $OA = a$ is the semi-major axis, and $OZ = b$, the semi-minor axis. The line QD, which is perpendicular to OA, is the directrix; F is focus; and FG, which is parallel to QD, is known as the semi-latus rectum (designated by l).

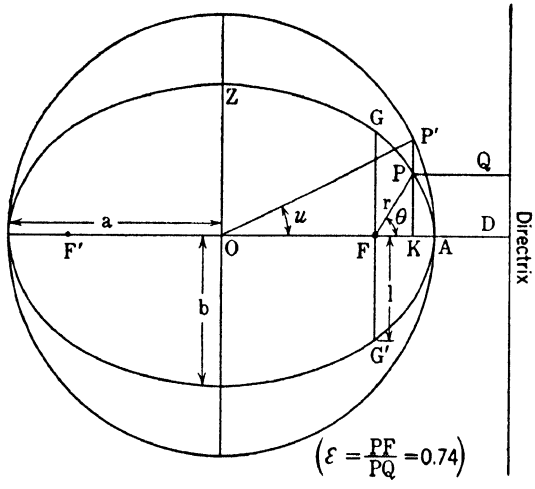


FIG. 23. Illustrating derivation of equation for ellipse in polar coördinates.

Let P be any point on the ellipse, and PQ the distance from QD. Let $FD = d$.

By the definition of an ellipse, the eccentricity is given by the relation

$$\epsilon = \frac{PF}{PQ} = \frac{r}{d - r\cos\theta}. \tag{i}$$

That is,

$$\frac{1}{r} = \frac{1}{\epsilon d} + \frac{\cos\theta}{d}. \tag{ii}$$

For $r = l$, $\cos\theta = 0$, so that $l = \epsilon d$. Therefore, equation (ii) becomes

$$\frac{l}{r} = 1 + \epsilon\cos\theta. \tag{iii}$$

Let r_0 and r_m designate minimum and maximum values of r, respectively.

For

$$r = FA = r_0, \quad \cos\theta = 1 \quad \text{and} \quad r_0 = \frac{l}{1+\epsilon} \tag{iv}$$

For

$$r = FO + a = r_m, \quad \cos\theta = -1 \quad \text{and} \quad r_m = \frac{l}{1-\epsilon} \tag{v}$$

Hence,

$$2a = r_0 + r_m = \frac{2l}{1-\epsilon^2}. \tag{vi}$$

Again

$$OF = \frac{l}{1-\epsilon} - a$$

$$= \frac{a(1-\epsilon^2)}{1-\epsilon} - a = a\epsilon.$$

Furthermore,

$$b^2 = (ZF)^2 - a^2\epsilon^2 = \epsilon^2\left(a\epsilon + \frac{l}{\epsilon}\right)^2 - a^2\epsilon^2$$

$$= a^2(1-\epsilon^2).$$

That is,

$$\frac{b}{a} = \sqrt{1-\epsilon^2}. \tag{vii}$$

The area of the ellipse is calculated most readily by using the equation in rectangular coördinates

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{viii}$$

$$\text{Area} = 4\int_a^0 y\,dx = 4\int_a^0 b\sqrt{1-\frac{x^2}{a^2}}\cdot dx.$$

Let

$$\frac{x}{a} = \sin\theta; \quad dx = a\cos\theta\,d\theta$$

$$\text{Area} = 4ab\int_0^{\frac{\pi}{2}} \cos^2\theta\,d\theta = \pi ab. \tag{ix}$$

In calculating some of the integrals which occur in the investigations on central orbits it has been found convenient to introduce a new variable u, designated the eccentric anomaly. In Fig. 23 let P' be the point at which the perpendicular through P on the major axis meets the circle with radius a (the semi-major axis) drawn about O as center. Then the angle $P'OK = u$.

Evidently,

$$FK = OK - OF = a \cos u - a\epsilon.$$

But from equation (iii)

$$FK = r \cos \theta = \frac{l - r}{\epsilon}.$$

Hence,

$$l - r = a\epsilon(\cos u - \epsilon).$$

Introducing equation (vi) it follows that

$$r = a(1 - \epsilon \cos u). \tag{x}$$

SUPPLEMENTARY NOTE 2

RELATION BETWEEN AVERAGE KINETIC AND POTENTIAL ENERGIES FOR CENTRAL ORBITS (VIRIAL THEOREM)

The following proof is given by J. H. Van Vleck.[18] Equations (4.3) and (4.6) lead to the relations of the form

$$\mu_i \ddot{x}_i = -\frac{\partial V}{\partial x_i}.$$

Multiplying both sides by $-x_i$, proceeding similarly with the equations corresponding to y_i and z_i, and summing over all the n particles of the system, we have

$$-\sum \mu_i (x_i \ddot{x}_i + y_i \ddot{y}_i + z_i \ddot{z}_i) = \sum \left(x_i \frac{\partial V}{\partial x_i} + y_i \frac{\partial V}{\partial y_i} + z_i \frac{\partial V}{\partial z_i} \right). \tag{i}$$

Now

$$\frac{d(x_i \dot{x}_i)}{dt} = x_i \ddot{x}_i + \dot{x}_i^2 \tag{ii}$$

and so forth, and furthermore $d(x_i \dot{x}_i)/dt$ is equal to zero on the average over a very long time interval if the separating distances and velocities of the particles never become infinite, and if the center of gravity is assumed at rest. Therefore, the average of the left-hand side of (i) equals that of the expression $\sum \mu_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2)$ which is twice the mean kinetic energy $\overline{T}$. [See equation (4.27).] Let the potential V be a homogeneous function of degree $n + 1$ in the coördinates. This will be the case if the forces of mutual action between particles vary as the nth power of the distance separating them. Then the expression on the right-hand side of (i) is by Euler's theorem on homogeneous functions [see equation (4.28)] equal to $(n + 1)V$. Consequently *the desired relation* is

$$2\overline{T} = (n + 1)\overline{V}. \tag{iii}$$

For *motion under the inverse square law*, $n = -2$, and consequently

$$2\overline{T} = -\overline{V}. \tag{iv}$$

It follows that for central orbits of this type

$$E = T + V = \frac{\overline{V}}{2}. \tag{v}$$

Since V is *negative* for all values of $r < \infty$, it follows that E also is negative.

[18] "Quantum Principles," pp. 20–21.

COLLATERAL READING

1. PAULING, L., and E. B. WILSON, JR., "Introduction to Quantum Mechanics." Chapter 1 gives a summary of the subject.
2. PAGE, L., "Introduction to Theoretical Physics," D. Van Nostrand Co., New York, Second Edition, 1936. The student should consult Chapter IV, "Advanced Dynamics," and Chapter XV, "Origin of Spectra." The method used for deducing Lagrange's equations and Hamilton's Principle depends upon the application of a simple theorem in the calculus of variation which the author derives in the first section of Chapter IV.
3. JOOS, G., "Theoretical Physics," Translated by I. M. Freeman, G. E. Stechert & Co., New York, 1934. Another invaluable work of reference on a large number of topics, including a section on classical dynamics and on wave mechanics. The method given in this treatise for deriving the Lagrange equations without the application of calculus of variation has been used in the preceding chapter.
4. VAN VLECK, J. H., "Principles of Quantum Mechanics," National Research Council Bull., 1926. A very comprehensive discussion of the subject of classical atomic mechanics.
5. LINDSAY, R. B., and MARGENAU, H., "Foundations of Physics," Chapter 3. The section on classical mechanics, like all the other sections of the book, should be consulted for a deeper understanding of the significance and importance of the fundamental concepts of dynamics.
6. BYERLY, W. E., "An Introduction to the Use of Generalized Coördinates in Mechanics and Physics," Ginn & Co., Boston, 1916. Like all this author's mathematical treatises, the topics are presented clearly, although, again, from a different point of view from that adopted in the preceding chapter.
7. TOLMAN, R. C., "Statistical Mechanics with Applications to Physics and Chemistry," Chemical Catalog Co., New York, 1927. Chapter II gives a concise presentation of classical mechanics.
8. WILSON, W., "Theoretical Physics," Vol. I, E. P. Dutton & Co., New York, 1931. Also a useful work of reference on classical mechanics. The student will find, in the chapter "Principles of Dynamics," another interesting method of deducing the fundamental relations.
9. BORN, M., "Vorlesungen über Atommechanik," Vol. I, Julius Springer, Berlin, 1925. A treatment of the classical dynamics and its application to atomic problems from the point of view of the Bohr theory.
10. OSGOOD, W. F., "Mechanics," Macmillan Co., New York, 1937. Contains a very comprehensive treatment of the subject, with no reference to atomic mechanics.
11. WEBSTER, A. G., "The Dynamics of Particles and of Rigid, Elastic and Fluid Bodies," G. E. Stechert, New York. This classical treatise is an extremely valuable work of reference.
12. WHITTAKER, E. T., "A Treatise on the Analytical Dynamics of Particles and Rigid Bodies," Macmillan Co., New York. For reference by the advanced student.
13. SLATER, J. C., and FRANK, N. H., "Introduction to Theoretical Physics," McGraw-Hill Book Co., New York, 1933.

CHAPTER V

THE LINEAR HARMONIC OSCILLATOR

5.1 Formulation of the S. Equation and Solution. The motion of a particle of mass μ acted upon by a restoring force proportional to the displacement is defined by the differential equation (2.6) which has the form

$$\mu \frac{d^2x}{dt^2} + kx = 0, \tag{1}$$

where k is a constant. The general solution of this equation, as shown in Chapter II, section 2, has the form

$$x = \sqrt{A^2 + B^2} \cdot \sin(\omega t + \delta), \tag{2}$$

where $\omega = \sqrt{k/\mu} = 2\pi\nu_0$, and ν_0 denotes the frequency of vibration, while δ is the phase angle.

In order to describe the behavior of such a linear harmonic oscillator from the point of view of wave mechanics, it is necessary, as a first step, to express the total energy E in the Hamiltonian form, that is, as a function $H(p, q)$ of the coördinates and corresponding momenta. As shown in equation (2.21), this relation has the form

$$E = H(p, q) = \frac{p^2}{2\mu} + \frac{kq^2}{2}, \tag{3}$$

where $p = \mu\,(dx/dt)$ = momentum, and q is used instead of x to designate the coördinate of position. Introducing the frequency of vibration $\nu_0 = \sqrt{k/\mu}/2\pi$, we derive the relation

$$E - U = E - \frac{kq^2}{2} = E - 2\pi^2\nu_0^2\mu q^2. \tag{4}$$

From equation (2.54) it follows that the corresponding S. equation has the form

$$\frac{d^2\phi}{dq^2} + \frac{8\pi^2\mu}{h^2}(E - 2\pi^2\nu_0^2\mu q^2)\phi = 0. \tag{5}$$

Let
$$a = \frac{8\pi^2\mu E}{h^2}, \tag{6}$$

and
$$b = \frac{4\pi^2\mu\nu_0}{h}. \tag{7}$$

Hence, we can replace (5) by the equation

$$\frac{d^2\phi}{dq^2} + (a - b^2q^2)\phi = 0. \tag{8}$$

Let us introduce the new variable

$$x = q\sqrt{b} = q \cdot 2\pi\frac{\sqrt{\mu h\nu_0}}{h}. \tag{9}$$

Since $\mu h\nu_0$ has the same dimensions as μE in the expression $\lambda = h/\sqrt{2\mu E}$, for the de Broglie wave length of a particle having kinetic energy E, it follows that $\sqrt{b}$ has the dimensions of a reciprocal length, and therefore x must be a *dimensionless* quantity, that is, a pure number.

Also from (9) it is seen that

$$\frac{d}{dq} = \frac{\sqrt{b} \cdot d}{dx}$$

and
$$\frac{d^2}{dq^2} = \frac{bd^2}{dx^2}.$$

Hence, the S. equation (8) becomes

$$\frac{d^2\phi}{dx^2} + \left(\frac{a}{b} - x^2\right)\phi = 0. \tag{10}$$

Now in order that ϕ shall have physical significance, it is necessary to obtain such solutions of equation (10) as will permit ϕ to be *finite and continuous for all values of* x *ranging from* $-\infty$ *to* $+\infty$. While in the case of a linear harmonic oscillator, in ordinary mechanics, the particle may vibrate only between the limits $x = \pm\sqrt{a/b}$ (since at these limits the value of U, the potential energy, becomes equal to the total energy E), there are no such limitations to the range of values of x in which ϕ may be finite. For, according to the Principle of Indeterminacy, there is no correlation between simultaneous values of position and velocity; the particle may be anywhere along the coördinate axis x. But it is obvious that $\phi\bar{\phi}$, the probability of occurrence of the particle,

as a function of x, must decrease continuously to zero as the values $x = \pm \infty$ are approached. It is the introduction of these " boundary conditions " that is essential for the solution of equation (10).

It will be observed that for values of $x \gg \sqrt{a/b}$, equation (10) assumes the form

$$\frac{d^2\phi}{dx^2} - x^2\phi = 0.$$

Writing this in the " operator " form,

$$(D^2 - x^2)\phi = (D + x)\,(D - x)\phi = 0,$$

it is seen that this equation is equivalent to the two *first*-order equations,

$$\frac{d\phi}{\phi} = \pm x dx,$$

that is,

$$d(\ln \phi) = \pm d\frac{x^2}{2},$$

where ln denotes logarithm to base ϵ.

While this solution is only approximate, since the factorization of the operator $(D^2 - x^2)$ is not justifiable for the case in which x is a variable, it does indicate that the solution of equation (11) should involve $\epsilon^{-x^2/2}$ as a factor. (Evidently a solution involving $\epsilon^{x^2/2}$ would not be reasonable, since it would correspond to an infinite value of $\phi(x)$ for $x \to \infty$.)

In consequence of these considerations we shall postulate a solution for equation (10) of the form

$$\phi(x) = \epsilon^{-\frac{x^2}{2}} \cdot \psi(x), \tag{11}$$

where $\psi(x)$ is a function of x, the nature of which must be determined from further considerations.

From equation (11) it follows that

$$\frac{d\phi}{dx} = -x\epsilon^{-\frac{x^2}{2}} \cdot \psi(x) + \epsilon^{-\frac{x^2}{2}} \cdot \frac{d\psi}{dx}$$

$$= -x\phi + \epsilon^{-\frac{x^2}{2}} \cdot \frac{d\psi}{dx},$$

while $$\frac{d^2\phi}{dx^2} = -\phi - x \cdot \frac{d\phi}{dx} - x\epsilon^{-\frac{x^2}{2}} \cdot \frac{d\psi}{dx} + \epsilon^{-\frac{x^2}{2}} \cdot \frac{d^2\psi}{dx^2}$$

$$= -\epsilon^{-\frac{x^2}{2}} \cdot \psi + x^2\epsilon^{-\frac{x^2}{2}} \cdot \psi - 2x\epsilon^{-\frac{x^2}{2}} \cdot \frac{d\psi}{dx} + \epsilon^{-\frac{x^2}{2}} \cdot \frac{d^2\psi}{dx^2}$$

$$= \epsilon^{-\frac{x^2}{2}} \left\{ \frac{d^2\psi}{dx^2} - 2x \cdot \frac{d\psi}{dx} + (x^2 - 1)\psi \right\}.$$

Since $\epsilon^{-x^2/2}$ is not equal to zero (except for $x = \pm\infty$), we obtain, by substitution of the value for $d^2\phi/dx^2$ indicated by (10) and (11), the equation to be solved in the form

$$\frac{d^2\psi}{dx^2} - 2x \cdot \frac{d\psi}{dx} + \left(\frac{a}{b} - 1\right)\psi = 0. \tag{12}$$

Let us express $\psi(x)$ in the form of a *polynomial*,

$$\psi(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0. \tag{13}$$

If $\phi(x)$, as defined by equation (11), is to vanish for $x = \pm\infty$, it follows that $x^n\epsilon^{-x^2/2}$ must also approach zero for large values of x. As x becomes infinite, x^n, as well as $\epsilon^{x^2/2}$, tends to become infinitely large, so that the value of the product $x^n\epsilon^{-x^2/2}$ is indeterminate. If, however, we make *n finite*, that is, let the series in (13) have *no powers greater than* x^n, then it can be shown that $x^n\epsilon^{-x^2/2}$ will always decrease to zero as x approaches $\pm\infty$.

From equation (13) we obtain the relations

$$2x \cdot \frac{d\psi}{dx} = 2na_n x^n + 2(n-1)a_{n-1}x^{n-1} + 2(n-2)a_{n-2}x^{n-2} + \ldots$$

$$\frac{d^2\psi}{dx^2} = n(n-1)a_n x^{n-2} + (n-1)\,(n-2)x^{n-3} + \ldots$$

If we substitute these equations in (12) it follows, since this equation is valid for any value of x, that the coefficient of each power of x must vanish identically. This leads to the " recursion formula "

$$(n+2)\,(n+1)a_{n+2} + \left(\frac{a}{b} - 1 - 2n\right)a_n = 0. \tag{14}$$

The series will end with x^n, if the coefficients a_{n+2}, a_{n+4}, etc., are each equal to zero, that is, if

$$\frac{a}{b} = 2n + 1, \tag{15}$$

where $n = 0, 1, 2$, etc.

It is evident that this deduction is analogous to that which was previously derived for an electron in a "box." The fact that (a/b) may assume only those values corresponding to the series of odd integers 1, 3, 5, etc., indicates that the S. equation (10) can have solutions which are physically significant *only for a series of discrete values of the energy E.* These values constitute the *characteristic energy values,* corresponding to a series of energy levels of the linear harmonic oscillator, and the corresponding solutions for ϕ constitute a series of characteristic functions.

Substituting for (a/b) in equations (6) and (7) it is found that the energy of the oscillator may assume any one of the series of discrete values defined by the relation

$$E_n = h\nu_0(n + \tfrac{1}{2}), \tag{16}$$

where E_n is therefore the *eigenvalue* corresponding to the *eigenfunction* ϕ_n.

On the basis of the older quantum theory, the energy states of the linear harmonic oscillator were deduced thus: According to equations (2.24) and (2.25)

$$p = \sqrt{2\mu E} \cdot \cos \omega t$$

$$q = \sqrt{\frac{2E}{k}} \cdot \sin \omega t.$$

Hence the action integral is given by

$$J = \oint pdq = 4E\sqrt{\frac{\mu}{k}} \int_{-\pi}^{\pi} \cos^2 \omega t \cdot d(\omega t) = \frac{4E}{\omega} \cdot \frac{\pi}{2} = \frac{E}{\nu_0}.$$

But $J = nh$ (quantum condition).

Therefore $E = nh\nu_0$.

On the other hand the S. solution leads to an additional term in the energy, of value $h\nu_0/2$. From (16) it also follows that the frequency in the nth state is

$$\nu_n = \nu_0(n + \tfrac{1}{2}).$$

This result is in better agreement with spectroscopic observations than that derived from the older theory. Furthermore, equation (16)

shows that the energy of a linear oscillator does not vanish at the absolute zero, but becomes equal to $h\nu_0/2$, thus solving definitely the problem of the existence of a *Nullpunktsenergie*.

We shall now proceed with the determination of the corresponding eigenfunctions. From equations (14) and (15) it follows that

$$\begin{aligned} n(n-1)a_n &= -[2n - 2(n-2)]a_{n-2} \\ &= 2 \cdot 2a_{n-2}, \end{aligned}$$

and that

$$\begin{aligned} (n-2)\,(n-3)a_{n-2} &= -[2n - 2(n-4)]a_{n-4} \\ &= -2 \cdot 2^2 a_{n-4}, \end{aligned}$$

whence we obtain the expression for ψ_n

$$\psi_n = a_n\left[x^n - \frac{n(n-1)x^{n-2}}{1 \cdot 2^2} + \frac{n(n-1)\,(n-2)\,(n-3)x^{n-4}}{1 \cdot 2 \cdot 2^4} \cdots\right], \tag{17}$$

where a_n is an arbitrary constant.

Now in investigating various types of functions, Hermite, a note mathematician of the nineteenth century, discovered a set of function known as "Hermitian" (or Hermite) polynomials, which are define by the relation[1]

$$H_n(x) = (-1)^n \epsilon^{x^2} \cdot \frac{d^n(\epsilon^{-x^2})}{dx^n}, \tag{18}$$

where $n = 0, 1, 2$, etc.

The expression defined by this relation is known as the nth Hermitia polynomial and is identical with (17) if we put $a_n = 2^n$. The first fi members of the series are readily determined.

$$H_0(x) = (-1)^0 \epsilon^{x^2} \cdot \epsilon^{-x^2} = 1.$$

$$H_1(x) = (-1)\epsilon^{x^2} \cdot \frac{d(\epsilon^{-x^2})}{dx} = 2x.$$

$$H_2(x) = (-1)^2 \epsilon^{x^2} \cdot \frac{d}{dx}(-2x \cdot \epsilon^{-x^2}) = 4x^2 - 2.$$

$$\begin{aligned} H_3(x) &= (-1)^3 \epsilon^{x^2} \cdot \frac{d}{dx}(4x^2\epsilon^{-x^2} - 2\epsilon^{-x^2}) \\ &= 8x^3 - 12x. \end{aligned}$$

$$H_4(x) = 16x^4 - 48x^2 + 12.$$

[1] These functions and their properties are described more fully in the mat matical treatises of Courant-Hilbert and Frank.-v. Mises.

From equation (18) the following relations may be derived.[2]

$$\frac{dH_n(x)}{dx} = H_n'(x) = 2nH_{n-1}(x). \tag{19}$$

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0. \tag{20}$$

The last relation makes it possible to derive the higher members of the series from the lower members. Combining equations (19) and (20), it follows that

$$H_{n+1}(x) - 2xH_n(x) + H_n'(x) = 0;$$

and differentiating with respect to x,

$$H_{n+1}'(x) - 2H_n(x) - 2xH_n'(x) + H_n''(x) = 0.$$

Hence,

$$2(n+1)H_n(x) - 2H_n(x) - 2xH_n'(x) + H_n''(x) = 0.$$

That is,

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0.$$

Evidently the last equation is identical with (12), if we put $(a/b) - 1 = 2n$.

Thus the eigenfunctions which satisfy the S. equation (10) are of the form

$$\phi_n = \epsilon^{-\frac{x^2}{2}} \cdot H_n(x).$$

We now have to consider the physical interpretation of these functions. As mentioned previously, $\phi_n\bar{\phi}_n dx$ or $\phi_n^2 dx$ (since in the present case ϕ_n is a *real function*) defines the probability of occurrence of the particle in the element of distance dx at the point x. It follows that

$$\int_{x_1}^{x_2} \phi^2 dx$$

is a measure of the probability of occurrence of the particle in the region $x_2 > x > x_1$, and since the oscillating particle must certainly be located in the range $x = \pm\infty$, it is necessary that we introduce *the normalizing factor* N_n, defined by the relation

$$\frac{1}{N_n^2}\int_{-\infty}^{\infty} \phi_n^2(x)dx = 1.$$

In order to illustrate the method used in determining the value of N_n and also for the purpose of demonstrating the *orthogonal* properties of

[2] Pauling and Wilson, "Introduction to Quantum Mechanics," pp. 77–82 give complete details of the derivation.

the eigenfunctions which satisfy the S. equation (10), it is necessary to introduce the values of certain definite integrals.

In any standard table of integrals[3] it is shown that

$$\int_0^\infty \epsilon^{-x^2} \cdot x^{2n+1}\, dx = \frac{n!}{2}. \tag{21}$$

$$\int_0^\infty \epsilon^{-x^2} \cdot x^{2n} dx = \frac{1 \cdot 3 \ldots (2n-1)\sqrt{\pi}}{2^{n+1}}. \tag{22}$$

It will be observed that, in the first of these integrals, x^{2n+1} changes sign with change of sign in x. The function $\epsilon^{-x^2} \cdot x^{2n+1}$ is therefore of the type designated as " odd," and resembles the trigonometric function $\sin x$.

Thus the area under the plot of the function for the range $x = 0$ to $x = \infty$ is positive, while the area for the range $x = 0$ to $x = -\infty$ is negative, and consequently the total area for the range $-\infty$ to $+\infty$ is zero. That is

$$\int_{-\infty}^\infty \epsilon^{-x^2} \cdot x^{2n+1} dx = 0. \tag{23}$$

On the other hand, in the case of the second integral, the function $\epsilon^{-x^2} \cdot x^{2n}$ is always positive, whether x is positive or negative. The function is thus of the type designated as " even," and resembles in this respect the function $\cos x$. Consequently,

$$\int_{-\infty}^\infty \epsilon^{-x^2} \cdot x^{2n} dx = 2\int_0^\infty \epsilon^{-x^2} \cdot x^{2n} dx = \frac{1 \cdot 3 \ldots (2n-1)\sqrt{\pi}}{2^n}. \tag{24}$$

By means of this equation and the expressions for $H_n(x)$ we derive the following values:

$$I_0 = \int_{-\infty}^\infty \phi_0^2(x) dx = \int_{-\infty}^\infty \epsilon^{-x^2} \cdot dx = \sqrt{\pi} = N_0^2,$$

$$I_1 = \int_{-\infty}^\infty \phi_1^2(x) dx = 4\int_{-\infty}^\infty \epsilon^{-x^2} \cdot x^2 dx = 2\sqrt{\pi} = N_1^2,$$

$$I_2 = \int_{-\infty}^\infty \phi_2^2(x) dx = \int_{-\infty}^\infty \epsilon^{-x^2} (4x^2 - 2)^2 dx = 2^2 \cdot 2\sqrt{\pi} = N_2^2,$$

where N_0, N_1, N_2 denote the normalizing factors corresponding to $n = 0$, 1, and 2, respectively.

[3] For example, L. Silberstein's "Mathematical Tables," G. Bell & Sons, London. See also Appendix III.

More generally, if we express ϕ_n in the form

$$\phi_n = \epsilon^{-\frac{x^2}{2}}(a_n x^n + a_{n-2}x^{n-2} + \ldots),$$

where a_n, a_{n-2}, etc., are the coefficients in the power series of the nth Hermitian polynomial, it is evident that

$$I_n = \int_{-\infty}^{\infty} \phi_n^2(x)dx = \sum \int_{-\infty}^{\infty} \epsilon^{-x^2} a_r a_s x^{r+s} dx,$$

where $\sum$ denotes that the sum of the series of terms is to be taken for the range of values $r + s = 2n$, $2n - 2$, and so forth. Hence, $r + s$ is always even, and it follows that I_n must always have a definite value

$$I_n = N_n^2$$

The value of N_n may be derived as follows. By definition,

$$N_n^2 = \int_{-\infty}^{\infty} \epsilon^{-x^2} \cdot H_n^2 dx,$$

where H_n is written instead of $H_n(x)$. From equation (18) it follows that

$$N_n^2 = (-1)^n \int_{-\infty}^{\infty} H_n \cdot \frac{d^n(\epsilon^{-x^2})}{dx^n} \cdot dx.$$

Now let us put

$$\frac{d^{n-1}(\epsilon^{-x^2})}{dx^{n-1}} = u,$$

and consider the result of differentiating the product $H_n u$. Since

$$\frac{d}{dx}(H_n u) = H_n u' + H_n' u,$$

we obtain by integration the relation

$$H_n u \Big]_{-\infty}^{\infty} = 0 = \int_{-\infty}^{\infty} H_n u' dx + \int_{-\infty}^{\infty} H_n' u dx.$$

That the first expression on the left vanishes is due to the fact that the presence of the term ϵ^{-x^2} in all the derivatives of this function makes u (and consequently uH_n) vanish at $x = \pm\infty$.

Combining the last relation with equation (19) it follows that

$$\int_{-\infty}^{\infty} H_n u' dx = -2n \int_{-\infty}^{\infty} H_{n-1} u dx.$$

That is,

$$(-1)^{-n}N_n^2 = \int_{-\infty}^{\infty} H_n \cdot \frac{d^n(\epsilon^{-x^2})}{dx^n}\,dx$$

$$= -2n\int_{-\infty}^{\infty} H_{n-1} \cdot \frac{d^{n-1}(\epsilon^{-x^2})}{dx^{n-1}}\,dx$$

$$= (-1)^2 \cdot 2^2 \cdot n(n-1)\int_{-\infty}^{\infty} H_{n-2} \cdot \frac{d^{n-2}(\epsilon^{-x^2})}{dx^{n-2}}\,dx,$$

and by continuing the same procedure, we finally deduce the relation

$$(-1)^{-n}N_n^2 = (-1)^n \cdot 2^n(n!)\int_{-\infty}^{\infty} H_0\epsilon^{-x^2}dx.$$

That is,

$$N_n^2 = (-1)^{2n}2^n(n!)\int_{-\infty}^{\infty} \epsilon^{-x^2}dx$$

$$= 2^n \cdot (n!) \cdot \pi^{\frac{1}{2}}. \tag{25}$$

Figure 24 shows plots of the normalized eigenfunctions, that is, of ϕ_n/N_n, for the values $n = 0, 1, 2, 3$, and 4. It will be observed that, for $n = 0, 2$, and 4, the curves are symmetrical with respect to a change

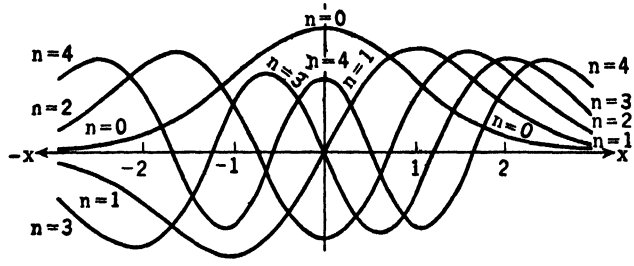


FIG. 24. Eigenfunctions for lower energy states of linear oscillator.

from positive to negative values of x (*even* functions), while for $n = 1$ and 3, the curves are antisymmetric with respect to such a change in sign of x. Excluding the nodes at $x = \pm\infty$, the number of these is always equal to n. This arises from the fact that $H_n(x)$ is a polynomial in x^n, and may therefore be expressed in the form $(x - a_1)(x - a_2) \ldots (x - a_n)$. Consequently $H_n(x)$ must have n roots, that is, n points along the axis of x at which $H_n(x) = 0$.

We shall now consider the integral

$$I_{nm} = \int_{-\infty}^{\infty} \phi_n\phi_m dx = \int_{-\infty}^{\infty} \epsilon^{-x^2}H_nH_m dx,$$

where *m is not equal to n.*

Assume $n = m + p$. Then we can proceed to calculate I_{nm} as in the case $m = n$, and we shall obtain the result

$$I_{nm} = \int_{-\infty}^{\infty} \phi_{m+p}\phi_m dx = \int \epsilon^{-x^2} H_{m+p} H_m dx$$
$$= K \int_{-\infty}^{\infty} H_p \epsilon^{-x^2} dx,$$

where K is a factor involving 2^p and the product $n(n-1) \ldots (n-m)$. Now from equation (18) it follows that

$$\int_{-\infty}^{\infty} H_p \epsilon^{-x^2} dx = \int_{-\infty}^{\infty} \frac{d^p(\epsilon^{-x^2})}{dx^p} dx,$$

and it is evident that the differential coefficient $d^p(\epsilon^{-x^2})/dx^p$ will consist of a series of terms of the form $a_r x^r \epsilon^{-x^2}$ where r has the maximum value p, and the signs of the coefficients will be alternately plus and minus. If p is odd, it is evident from equation (23) that the integral I_{nm} will be equal to zero. If p is even, it is not self-evident that the same conclusion will be valid, and a more elaborate proof is required. However, the result may be demonstrated by an actual calculation for a simple case.

Thus, let us consider the product $\phi_1\phi_3$. From the values of the corresponding Hermitian polynomials, it is seen that

$$\int_{-\infty}^{\infty} \phi_1\phi_3 dx = 8 \int_{-\infty}^{\infty} \epsilon^{-x^2}(2x^4 - 3x^2) dx$$
$$= \frac{16 \cdot 3\sqrt{\pi}}{2^2} - \frac{24\sqrt{\pi}}{2} = 0.$$

Similarly,

$$\int_{-\infty}^{\infty} \phi_2\phi_4 dx = \int_{-\infty}^{\infty} \epsilon^{-x^2}(4x^2 - 2)\,(16x^4 - 48x^2 + 12) dx$$
$$= \int_{-\infty}^{\infty} (64\epsilon^{-x^2} \cdot x^2 - 224\epsilon^{-x^2} \cdot x^2 + 144\epsilon^{-x^2} \cdot x^2 - 24\epsilon^{-x^2}) dx$$
$$= \frac{64 \cdot 1 \cdot 3 \cdot 5}{2^3} - \frac{224 \cdot 1 \cdot 3}{2^2} + \frac{144 \cdot 1}{2} - 24 = 0.$$

That is, the value of $\phi_n\phi_m$ alternates between positive and negative values, as may be seen by inspecting the curves shown in Fig. 24, with the result that $I_{nm} = 0$. This deduction may be stated in the generalized form that the *solutions of the S. equation form an orthogonal set of eigenfunctions.* As mentioned in Chapter III this conclusion is valid for the solutions obtained for the S. equation in all cases, and may be demon-

strated to be a logical consequence of a very fundamental mathematical theorem.[4] As shown in subsequent sections, this property of the solutions of the S. equation is of extremely great importance in dealing with a number of problems which arise in quantum mechanics.

5.2 Distribution Functions for Harmonic Oscillator. We shall now consider further the interpretation, on the basis of wave mechanics, of the behavior of a particle which, according to classical mechanics, executes harmonic vibrations with frequency ν_0.

From equation (9) it is observed that in terms of q, the actual displacement, the dimensionless variable is defined by the relation

$$x = q\sqrt{b} = 2\pi q\sqrt{\frac{\mu\nu_0}{h}}.$$

In classical mechanics, the maximum amplitude of oscillation is given by

$$q_0 = \sqrt{\frac{E}{2\pi^2\mu\nu_0^2}},$$

as is evident from equation (4), since, for this value of q, $E - U = 0$. Substituting for E from equation (16) and using equation (9), the corresponding maximum value of x is found to be

$$x_0 = \sqrt{\frac{a}{b}} = \sqrt{2n+1}. \tag{26}$$

Hence, the motion of the oscillating particle, from the classical mechanics point of view, would be given by the relation

$$x = \sqrt{2n+1} \cdot \sin(2\pi\nu_n t), \tag{27}$$

showing that x oscillates between $+\sqrt{2n+1}$ and $-\sqrt{2n+1}$.

From this it is possible to calculate the probability Pdx that the particle will be found between x and $x + dx$.[5] Since Pdx is proportional to the element of time dt required for the particle to pass from x to $x + dx$,

$$Pdx = Adt,$$

where A is a constant which satisfies the condition that

$$\int_{-x_0}^{x_0} Pdx = 1. \tag{28}$$

[4] A proof for this statement is given in the supplementary note 1 for the case in which the S. equation involves only one coördinate variable.

[5] The following remarks constitute an amplification of the calculation given by Condon and Morse, "Quantum Mechanics," p. 51.

From equation (27) it follows that

$$dx = \sqrt{2n+1} \cdot 2\pi\nu_n \cos(2\pi\nu_n t)dt$$
$$= 2\pi\nu_n(\sqrt{2n+1-x^2})dt.$$

Hence,

$$Adt = \frac{Adx}{2\pi\nu_n\sqrt{2n+1-x^2}}.$$

Introducing equation (28), according to which

$$\int_{-\sqrt{2n+1}}^{\sqrt{2n+1}} \frac{Adx}{2\pi\nu_n\sqrt{2n+1-x^2}} = 1,$$

it is found that

$$A = 2\nu_n,$$

and that therefore

$$Pdx = \frac{dx}{\pi\sqrt{2n+1-x^2}}. \tag{29}$$

This result shows that, for the classical oscillator, P increases from $1/(\pi\sqrt{2n+1})$ at $x = 0$ to ∞ at $x = x_0 = \pm\sqrt{2n+1}$. On the other hand, S.'s solution leads to finite values of ϕ_n even for values of $|x| > |x_0|$. That is, ϕ_n^2 *has a definite value even for values of* x (*or* q) *which are forbidden by classical mechanics.*

In the region outside the classical limits, the potential energy U is greater than the total energy E, and hence the de Broglie wave length is imaginary. From the S. equation (10) it follows that

$$\frac{d^2\phi}{dx^2} = -\left(\frac{a}{b} - x^2\right)\phi = -(2n+1-x^2)\phi.$$

Hence for $x = x_0$ $d^2\phi/dx^2 = 0$, which shows that this constitutes a *point of inflection.* For values of $|x| > |x_0|$, ϕ decreases continuously without exhibiting any nodes, and, as mentioned previously, for very large values of x, ϕ varies as $\epsilon^{-x^2/2}$, that is, decreases rapidly without increase in x.

Plotting the values of ϕ_n^2/N_n^2 against x, a series of curves such as those shown in Fig. 25 is obtained, and designated by A. (The significance of $1/N_n^2$ consists in the fact that it makes the area under each of these curves equal to $\frac{1}{2}$. Since similar curves, symmetrical with respect to the ϕ^2-axis, may be plotted for negative values of x, the *total* area under the curve giving ϕ^2/N^2 as a function of *all* values of x is equal to 1.)

The curves designated by B in Fig. 25 give the probability distribution function as calculated from equation (29), that is, *according to classical mechanics.* The difference in results derived from the two points of view is due, as emphasized in the remarks on the " tunneling effect," to the complete lack of association in quantum mechanics between position and velocity. That is, we have here another illustration of the application of the Principle of Indeterminacy.

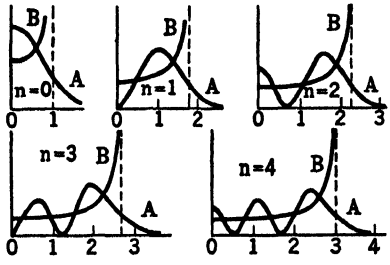


FIG. 25. Probability of occurrence of oscillating particle according to quantum mechanics (Curve A) and according to classical theory (Curve B).

That the probability for the occurrence of the oscillating particle in the region $|x| > |\sqrt{2n+1}|$ is considerable, is readily shown by calculating the value of $(2/N^2)\int_{x_0}^{\infty}\phi^2 dx$ in the case of the first characteristic function, $\phi_0 = \epsilon^{-x^2/2}$ corresponding to $E = h\nu_0/2$.

Here we find[6] (since $x_0 = 1$),

$$\frac{2}{\sqrt{\pi}}\int_1^{\infty}\epsilon^{-x^2}dx = 0.1573.$$

That is, there is a 15.7 per cent probability that the particle will occur in the region for which the potential energy exceeds the total energy.

Curves such as those shown in Fig. 25 are known as probability distribution curves. As emphasized previously, the Principle of Indeterminacy states that it is impossible to fix simultaneously the coördinates of position and the magnitudes of momenta. It follows logically that the solution of a S. equation can have only a statistical interpretation as regards each of the variables specifying position and momentum. The probability distribution curve represents the information about these variables thus derived.

It will be observed that, in the graph of any such distribution curve, the ordinate is really a differential coefficient. Thus, in Fig. 25, the ordinate P should be defined as

$$P_x = \left(\frac{d\rho}{dx}\right)_x,$$

where ρ is the area under the curve between $x = 0$ and x. In the present case this area gives the probability for the occurrence of the particle

[6] See supplementary note 2.

between 0 and x, where the total probability of occurrence from $x = -\infty$ to $x = +\infty$ is taken as unity.

In classical physics the best-known curve of a similar nature is that which represents the Maxwell-Boltzmann distribution law for the velocities of molecules, which has the form[7]

$$\frac{1}{N} \cdot \frac{dN}{dc} = 4\pi \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} \epsilon^{-\frac{mc^2}{2kT}} .$$

In this equation N designates the total number of molecules, c the random velocity, m the mass of the molecule, k the Boltzmann constant, and T the absolute temperature of the gas.

From this equation it is possible to derive values of the average velocity, the root-mean-square velocity, and the most probable velocity. Similar equations may be derived for the distribution of energy, and for distribution of velocities in any given direction.

Another illustration is the plot representing the distribution of energy with frequency in the radiation from a black body at any given temperature T.[8]

This relation has the form

$$\frac{dE}{d\nu} = E_\nu = \frac{2\pi\nu^2}{c^2} \cdot \frac{h\nu}{\epsilon^{\frac{h\nu}{kT}} - 1},$$

where ν = frequency, c = velocity of light, and h = Planck's constant.

The total energy radiated is given by

$$E = \int_0^\infty E_\nu d\nu,$$

that is, by the area under the curve which gives E_ν as a function of ν.

5.3 Determination of Average Values. A problem which is of much interest is the determination of the average values of the coördinates (designated by q) or of the corresponding momenta (designated by p). The relations used for calculating these values in quantum mechanics follow readily from equations of a similar nature which have been derived in ordinary mechanics for calculating the center of gravity of a body.

Thus, let us consider a rod of uniform cross section and given length, but of variable density ρ, so that $\rho = \rho(x)$ is a function of x, the dis-

[7] This is discussed very fully in Tolman's "Statistical Mechanics."

[8] See S. Dushman, Taylor's "Treatise," Chapter XVI.

tance from one end. Let x_0 denote the coördinate of the center of gravity. The moment of any element dx about the center of gravity is given by the product,

$$\text{mass of element} \times \text{distance} = \rho dx \cdot (x - x_0).$$

Now the center of gravity is that point about which the total moment of all the forces acting on each element is zero. Hence, if l = length of rod,

$$0 = \int_0^l \rho(x - x_0)dx,$$

or

$$x_0 = \frac{\int_0^l \rho x dx}{\int_0^l \rho dx}.$$

For example, if $\rho = \rho_0 x$, that is, if the linear density is proportional to the distance x from one end of the rod, then

$$x_0 = \frac{\rho_0 \int_0^l x^2 dx}{\rho_0 \int_0^l x dx} = \frac{2l}{3}.$$

Similarly, in the case of two coördinates of position, the coördinates x_0, y_0 of the center of gravity are given by the relations

$$x_0 = \frac{\int\int \rho x dx dy}{\int\int \rho dx dy}, \quad y_0 = \frac{\int\int \rho y dx dy}{\int\int \rho dx dy},$$

where $\rho = \rho(x, y)$, and the integration is carried out over the whole area of the surface under consideration. It is evident that the denominator always corresponds to the total mass. The extension of these equations to three dimensions is obvious.

Now, in quantum mechanics, the magnitude $\phi\bar{\phi}$, or ϕ^2 (in case of real functions), is a measure of the relative concentration or density at any point, per unit length, per unit area, or per unit volume (depending upon whether ϕ is a function of one, two, or three variables, respectively). Consequently, the average values of variables are given by formulas similar to those derived above, in which ρ is replaced by ϕ^2 (or $\phi\bar{\phi}$). If ϕ is a function of x only, the average values of x and p, the correspond-

ing momentum, are given by the relations

$$x_{\text{Av.}} = \frac{\int_{-\infty}^{\infty} \bar{\phi} x \phi dx}{\int_{-\infty}^{\infty} \bar{\phi} \phi dx}, \tag{30}$$

$$p_{\text{Av.}} = \frac{\int_{-\infty}^{\infty} \bar{\phi} p \phi dx}{\int_{-\infty}^{\infty} \bar{\phi} \phi dx}. \tag{31}$$

In the case of *normalized eigenfunctions, the denominator is equal to unity,* corresponding to the physical interpretation that the total concentration is to be taken as the standard of reference.

It will be observed that in equation (31) $\bar{\phi} p \phi$ has been written instead of $\bar{\phi} \phi p$. The reason for this is evident when we remember that in quantum mechanics p has the significance of the *operator,*

$$p = \frac{h}{2\pi i} \cdot \frac{\partial}{\partial x}.$$

(See Chapter II, section 7.)

Hence,

$$\bar{\phi} p \phi dx = \frac{h\bar{\phi}}{2\pi i} \cdot \frac{d\phi}{dx} \cdot dx,$$

while $\bar{\phi} \phi p$ has no physical significance. That is, p and ϕ are *non-commutative* symbols. On the other hand, x and ϕ are commutative, so that we could write $\bar{\phi} \phi x$, but it is customary to write the product in the form given in (30) in order to preserve the analogy with (31).

As an illustration of the application of these equations let us consider the following problem. Given the state of a system of particles represented by the eigenfunction

$$\phi = A\epsilon^{iax} + B\epsilon^{-iax}, \tag{32}$$

what are the average values of x and of the momentum?

The physical interpretation of the two terms on the right-hand side of this equation has been considered previously in Chapter II, so that the following remarks merely supplement the previous discussion.

Let us now apply equation (30) to each term in equation (32). For the first term

$$x_{\text{Av.}} = \frac{\int_{-\infty}^{\infty} A\bar{A} x dx}{\int_{-\infty}^{\infty} A\bar{A} dx} = \frac{A\bar{A}x^2\Big]_{-\infty}^{\infty}}{2A\bar{A}\int_{-\infty}^{\infty} dx} = 0,$$

which may also be deduced from the fact that x changes sign in passing from $x = -\infty$ through $x = 0$ to $x = +\infty$, and that therefore the area under the line $y = A\bar{A}x$ and the axis of x for $x = 0$ to $x = \infty$ cancels the area for $x = -\infty$ to $x = 0$.

Similarly, $x_{\text{Av.}}$ for $B\epsilon^{-i\alpha x}$ is found to be equal to zero. That is, the function ϕ in (32) represents a state in which there are just as many particles on one side of $x = 0$ as on the other.

For the term $\epsilon^{i\alpha x}$, the average momentum is given by

$$p_{\text{Av.}} = \frac{\int_{-\infty}^{\infty} \epsilon^{-i\alpha x} p \cdot \epsilon^{i\alpha x} dx}{\int_{-\infty}^{\infty} \epsilon^{-i\alpha x} \epsilon^{i\alpha x} dx}$$

$$= \frac{\frac{h}{2\pi} \int_{-\infty}^{\infty} \epsilon^{-i\alpha x} \alpha \epsilon^{i\alpha x} dx}{\int_{-\infty}^{\infty} dx}$$

$$= \frac{h\alpha}{2\pi}.$$

Similarly, for the term $\epsilon^{-i\alpha x}$, the average momentum is found to be $-h\alpha/2\pi$.

In the case of the function ϕ, defined by equation (32), the average momentum is obtained from the relation

$$p_{\text{Av.}} = \frac{c_1 \cdot p_1 + c_2 \cdot p_2}{c_1 + c_2},$$

where c_1 denotes the concentration per unit length of particles with momentum p_1, and c_2 the concentration of particles with momentum p_2. In the present case c_1 refers to the particles possessing momentum $h\alpha/(2\pi)$, and c_2 refers to those with momentum $-h\alpha/(2\pi)$. The values of the concentrations are given by

$$c_1 = A\epsilon^{i\alpha x} \cdot \bar{A}\epsilon^{-i\alpha x} = |A|^2; \quad c_2 = |B|^2.$$

Hence,

$$p_{\text{Av.}} = \frac{|A|^2 - |B|^2}{|A|^2 + |B|^2} \cdot \frac{h\alpha}{2\pi}.$$

From this result it is evident that the sign and magnitude of $p_{\text{Av.}}$ are governed by the relative magnitudes of $|A|^2$ and $|B|^2$.

As another illustration we shall determine average values of x and

of p for the linear harmonic oscillator. In terms of the normalized functions,

$$x_{\text{Av.}} = \frac{1}{N_n^2}\int_{-\infty}^{\infty} \epsilon^{-x^2} H_n^2(x) \cdot x dx.$$

Now $\epsilon^{-x^2}H_n^2$ is always positive, so that the product of this function with x changes sign at $x = 0$. Hence the integral vanishes. That is, there is an equal likelihood for the occurrence of the particle on either side of the origin.

Again,

$$p_{\text{Av.}} = \frac{1}{N_n^2}\int_{-\infty}^{\infty} \epsilon^{-\frac{x^2}{2}} H_n(x) \cdot \frac{h}{2\pi i} \cdot \frac{d}{dq}\left[\epsilon^{-\frac{x^2}{2}} H_n(x)\right] dx.$$

But

$$\frac{d}{dq} = \sqrt{b} \cdot \frac{d}{dx}.$$

Hence,

$$p_{\text{Av.}} = \frac{h\sqrt{b}}{2\pi i N_n^2}\int_{-\infty}^{\infty} \epsilon^{-\frac{x^2}{2}} H_n(x) \frac{d}{dx}\left[\epsilon^{-\frac{x^2}{2}} H_n(x)\right] dx.$$

Now

$$\frac{d}{dx}\left[\epsilon^{-\frac{x^2}{2}} H_n(x)\right] = H_n'(x) \cdot \epsilon^{-\frac{x^2}{2}} - x\epsilon^{-\frac{x^2}{2}} H_n(x),$$

so that the expression to be integrated is

$$\epsilon^{-x^2}[H_n(x) \cdot H_n'(x) - xH_n^2(x)].$$

It is evident that each of the terms in this expression is an odd function. For, if $H_n(x)$ is a polynomial of the form $a_k x^k$, $H_n'(x)$ must be of the form $a_k k x^{k-1}$. Consequently the product $H_n H_n'$ will consist of a series of odd powers of x, while xH_n^2 must be an odd function for the reason mentioned in discussing the value of $x_{\text{Av.}}$. Therefore $p_{\text{Av.}}$ must also be zero.

That is, *it is not possible to perform any experiment by which either the position or velocity of the oscillating particle may be determined at any given instant.*

The argument by which this conclusion was derived was essentially to the effect that in both cases the expressions to be integrated contained *odd* powers of x. It is therefore evident that the average values of q^2 and p^2 will not vanish.

It should be observed in this connection that the operator p refers to the coördinate of position q, so that

$$p = \frac{h}{2\pi i} \cdot \frac{d}{dq} = \frac{h\sqrt{b}}{2\pi i} \cdot \frac{d}{dx}$$

and

$$p^2 = -\frac{h^2 b}{4\pi^2} \cdot \frac{d^2}{dx^2}.$$

By definition of the average value, as given in equations (30) and (31),

$$x^2_{\text{Av.}} = \frac{1}{N_n^2} \int_{-\infty}^{\infty} \epsilon^{-x^2} \cdot x^2 H_n^2(x) dx, \tag{33}$$

$$p^2_{\text{Av.}} = -\frac{h^2 b}{4\pi^2 N_n^2} \int_{-\infty}^{\infty} \epsilon^{-\frac{x^2}{2}} \cdot H_n(x) \frac{d^2}{dx^2} \left\{ \epsilon^{-\frac{x^2}{2}} H_n(x) \right\} dx. \tag{34}$$

The exact calculation of these integrals in the general case involves more elaborate methods than those that may be discussed in the present connection. However, it is possible to carry out the calculation for the case $H_0(x)$, that is, for the zero energy state, for which the energy $E_0 = h\nu_0/2$ and $H_0(x) = 1$. Substituting in equations (33) and (34), we obtain from equation (24) the relations

$$x^2_{\text{Av.}} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \epsilon^{-x^2} \cdot x^2 dx = \frac{1}{2},$$

and

$$p^2_{\text{Av.}} = \frac{1}{\sqrt{\pi}} \cdot \frac{bh^2}{4\pi^2} \int_{-\infty}^{\infty} \epsilon^{-\frac{x^2}{2}} \left(\epsilon^{-\frac{x^2}{2}} - x^2 \epsilon^{-\frac{x^2}{2}} \right) dx$$

$$= \frac{bh^2}{4\pi^2} \cdot \frac{1}{\sqrt{\pi}} \left(\sqrt{\pi} - \frac{\sqrt{\pi}}{2} \right)$$

$$= \frac{bh^2}{8\pi^2} = \frac{h^2 \cdot 4\pi^2 \mu \nu_0}{h \cdot 8\pi^2} \text{ from (7)}$$

$$= \frac{\mu h \nu_0}{2} = \mu E_0. \tag{35}$$

According to classical mechanics, the kinetic energy of the oscillating particle is $T = p^2/(2\mu)$, and this varies from $T = 0$ at the limits of vibration to $T = E_0$ when passing through $q = 0$. Consequently, the average value of T is $E_0/2$, and as will be observed from equation (35), the average value of $p^2/(2\mu)$ on the basis of wave mechanics is found to be equal to $E_0/2$, as in the classical case.

Since $x^2_{\text{Av.}} = 0.5$, it follows that $\sqrt{x^2_{\text{Av.}}} = \pm\sqrt{0.5} = \pm 0.707$, where $|x_0| = 1$ represents the maximum amplitude of vibration.

5.4 Probabilities of Transition.[9] In connection with the calculation of *probabilities of transition* from one energy state to another, it is necessary to calculate the integral

$$J_{mn} = \left(\frac{1}{N_m N_n}\right)\int_{-\infty}^{\infty} \phi_n \phi_m x dx, \tag{36}$$

where n and m refer to two *different* states for which the energies are E_n and E_m respectively. The integral in (36) is known as a *matrix component* integral, because it is of fundamental importance in the type of quantum mechanics introduced by W. Heisenberg which has been designated as *matrix calculus*.

In terms of Hermitian polynomials

$$J_N = J_{mn} N_m N_n = \int_{-\infty}^{\infty} \epsilon^{-x^2} H_n(x) \cdot H_m(x) x dx. \tag{37}$$

Since

$$\frac{d}{dx}(\epsilon^{-x^2} H_n H_m) = -2x\epsilon^{-x^2} H_n H_m + \epsilon^{-x^2}(H_n H_m' + H_m H_n'),$$

we can, by repeated application of the standard formula for integration by parts and cancelation of identical terms of opposite sign, write

$$2J_N = -\epsilon^{-x^2} \cdot H_n H_m \Big]_{-\infty}^{\infty} + \int \epsilon^{-x^2} H_n H_m' dx + \int \epsilon^{-x^2} H_m H_n' dx.$$

(For convenience we shall omit the limits of integration in writing the integral.)

Since $\epsilon^{-x^2/2} H_n(x)$ vanishes at $x = \pm\infty$, the first term on the right-hand side disappears. Also, according to (19), we may replace H_m' and H_n', respectively, by the expressions $2mH_{m-1}$ and $2nH_{n-1}$. Hence we derive the relation

$$J_N = m\int \epsilon^{-x^2} H_n H_{m-1} dx + n\int \epsilon^{-x^2} H_m H_{n-1} dx. \tag{38}$$

[9] This topic is discussed in Chapter XV.

Now, as stated previously, the functions $\phi_n = \epsilon^{-x^2/2}H_n$ form an orthogonal system, so that for $n \neq k$,

$$\int \phi_n \phi_k dx = \int \epsilon^{-x^2} H_n H_k dx = 0.$$

Therefore, the integrals in (38) will each be equal to zero, *unless*

$$n = m - 1, \text{ that is } m = n + 1,$$

or

$$m = n - 1.$$

In other words, J_{mn} has a definite value *only if* m *differs from* n *by* $+1$ *or* -1. The physical interpretation of this conclusion is that in the case of the linear harmonic oscillator the *only possible transitions are those between* adjacent states, for which $\Delta n = \pm 1$, and therefore $\Delta E = \pm h\nu_0$.

Let us consider the case $m = n + 1$. Then

$$J_{mn} = \frac{(n+1)}{N_{n+1}N_n}\int \epsilon^{-x^2} H_n^2 dx = \frac{N_n(n+1)}{N_{n+1}}.$$

Using equation (25), we obtain the result

$$J_{n+1,n} = \sqrt{\frac{n+1}{2}}. \tag{39}$$

Similarly, it follows that

$$J_{n,n-1} = \sqrt{\frac{n}{2}}. \tag{40}$$

SUPPLEMENTARY NOTE 1

PROOF THAT SOLUTIONS OF THE SCHROEDINGER EQUATION FOR ONE DEGREE OF FREEDOM ARE ORTHOGONAL

For the case in which the S. equation is expressed in terms of a single coördinate variable, the proof that the solutions form an orthogonal system is as follows.

Let us consider the S. equation of the form

$$\frac{d^2\psi}{dx^2} + k^2(E - V)\psi = 0, \qquad \text{(i)}$$

where $k^2 = 8\pi^2\mu/h^2$.

Let ψ_m and ψ_n denote any two solutions valid for the range $x = -\infty$ to $x = +\infty$, and let E_m and E_n denote the corresponding eigenvalues, which are not identical. In the general case ψ_m and ψ_n will be complex functions, so that in order to obtain a *real magnitude* we must take the product $\psi_m\bar{\psi}_m$ or $\bar{\psi}_n\psi_n$.

In consequence of (i), we have the two relations

$$\frac{d^2\psi_m}{dx^2} + k^2(E_m - V)\psi_m = 0, \qquad \text{(ii)}$$

$$\frac{d^2\bar{\psi}_n}{dx^2} + k^2(E_n - V)\bar{\psi}_n = 0. \qquad \text{(iii)}$$

If we multiply the first of these equations by $\bar{\psi}_n$ and the second by ψ_m, and integrate between the limits $x = \pm\infty$, we obtain the relations

$$\int \bar{\psi}_n\psi''_m dx + \int k^2(E_m - V)\psi_m\bar{\psi}_n dx = 0, \qquad \text{(iv)}$$

$$\int \psi_m\bar{\psi}''_n dx + \int k^2(E_n - V)\psi_m\bar{\psi}_n dx = 0, \qquad \text{(v)}$$

where $\psi''_m = \frac{d^2\psi_m}{dx^2}$, and similarly for $\bar{\psi}''_n$.

Subtracting (iv) from (v), the term involving V cancels out, so that

$$\int (\psi_m\bar{\psi}''_n - \bar{\psi}_n\psi''_m)dx + k^2(E_n - E_m)\int \psi_m\bar{\psi}_n dx = 0. \qquad \text{(vi)}$$

Now

$$\frac{d}{dx}(\bar{\psi}_n\psi'_m) = \bar{\psi}_n\psi''_m + \bar{\psi}'_n\psi'_m.$$

Also

$$\frac{d}{dx}(\psi_m\bar{\psi}'_n) = \psi_m\bar{\psi}''_n + \bar{\psi}'_n\psi'_m.$$

Hence,

$$\int(\psi_m\bar{\psi}''_n - \bar{\psi}_n\psi''_m)dx = \int\frac{d}{dx}(\psi_m\bar{\psi}'_n - \bar{\psi}_n\psi'_m)dx = (\psi_m\bar{\psi}'_n - \bar{\psi}_n\psi'_m)\Big]_{-\infty}^{\infty} = 0,$$

since ψ_m and $\bar{\psi}_n$, as well as their differential coefficients, all tend to become equal to zero as the limits are approached.

Consequently, equation (vi) reduces to the relation

$$k^2(E_n - E_m)\int_{-\infty}^{\infty}\psi_m\bar{\psi}_n dx = 0. \qquad \text{(vii)}$$

Since k^2 is a finite quantity, and E_n is *not equal* to E_m, it follows that

$$\int_{-\infty}^{\infty}\psi_m\bar{\psi}_n dx = 0. \qquad \text{(viii)}$$

That is, *any two eigenfunctions corresponding to two different energy states are orthogonal.*

The use of the designation " orthogonal " is due to the analogy between relation (viii) and that which expresses the condition that two vectors shall be at right angles. Thus we can represent a momentum $p = \mu v$, in both magnitude and direction, by a straight line of length p (a scalar magnitude) which is drawn with respect to the three rectangular axes of coördinates in the direction of motion. The line thus drawn constitutes a vector. Let θ_x, θ_y, θ_z be the angles between this line and the three axes of coördinates. Then the components of momentum along the three axes are given by

$$p_x = p\cos\theta_x,$$
$$p_y = p\cos\theta_y,$$
$$p_z = p\cos\theta_z,$$

so that

$$p_x^2 + p_y^2 + p_z^2 = p^2.$$

Now suppose we have another vector P with components $P_x = P\cos\eta_x$, etc., where η_x is the angle between the vector P and the axis of x.

Let ϕ denote the angle between the two vectors p and P. Then, it follows that

$$\cos \phi = \cos \theta_x \cdot \cos \eta_x + \cos \theta_y \cdot \cos \eta_y + \cos \theta_z \cdot \cos \eta_z.$$

Consequently,

$$pP \cos \phi = p_x P_x + p_y P_y + p_z P_z.$$

The product on the left-hand side is a scalar magnitude. In vector notation this is written $\mathbf{p} \cdot \mathbf{P}$ and is known as the "dot product" of the vectors.

In order that the two vectors shall be at right angles, it is necessary that $\phi = \pi/2$, and that $\cos \phi = 0$. Hence, the condition for the orthogonality of the two vectors is given by the relation

$$p_x P_x + p_y P_y + p_z P_z = 0.$$

If now we assume a space of, say, n dimensions and consider two vectors in this space, each with n components along the n-axes of coördinates, the last equation becomes

$$\sum_1^n p_k P_k = 0,$$

where k has all values ranging from 1 to n, and $\sum$ denotes the sum of all these products. By extending this argument to a space of an infinite number of dimensions, and regarding ϕ_m and $\bar{\phi}_n$ as vectors in this space, equation (viii) follows logically.

SUPPLEMENTARY NOTE 2

THE GAUSS ERROR FUNCTION

The curve for $\phi_0(x) = \epsilon^{-x^2/2}$ is essentially the same as the Gauss error curve, which is ordinarily expressed in the form[10] $y = \frac{h}{\sqrt{\pi}} \cdot \epsilon^{-h^2x^2}$. In this expression h corresponds to the absolute measure of precision, and x is the magnitude of the deviation from the mean value. Thus y is a measure of the probability of occurrence of a deviation of magnitude x.

The *probability integral* is given by

$$P = \frac{2h}{\sqrt{\pi}} \int_0^x \epsilon^{-h^2x^2} dx.$$

For $h = 1$, the integral is known as Gauss's Error Integral, or the " erf " function, and is designated by

$$\theta(x) = \frac{2}{\sqrt{\pi}} \int_0^x \epsilon^{-x^2} dx.$$

The following table gives values of this integral for different values of x, as taken from L. Silberstein's " Mathematical Tables."

x	$\frac{2}{\sqrt{\pi}} \int_0^x \epsilon^{-x^2} dx$	$\frac{2}{\sqrt{\pi}} \int_x^\infty \epsilon^{-x^2} dx$
0.0	0.0000	1.0000
0.25	0.2763	0.7237
0.5	0.5205	0.4795
0.75	0.7112	0.2888
1.0	0.8427	0.1573
1.5	0.9661	0.0339
3.0	0.99998	0.00002

COLLATERAL READING

The problem of the linear harmonic oscillator is discussed in all the treatises on quantum mechanics (see list in Appendix I), but the reader will find it of special assistance to consult the following:

1. Condon, E. U., and Morse, P. M., "Quantum Mechanics," pp. 47–52.
2. Pauling, L., and Wilson, E. B., Jr., "Introduction to Quantum Mechanics," pp. 67–82.
3. Darrow, K. K., *Bell System Tech. J.*, **6**, 653 (1927).
4. Sommerfeld, A., "Wave Mechanics," Chapter I.

[10] See, for instance, Mellor's "Higher Mathematics."

CHAPTER VI

THE RIGID ROTATOR

6.1 The Schroedinger Equation in Three Dimensions. In order to discuss such problems as that of the rotational energy states of a diatomic molecule, or the energy states of a hydrogen-like atom, it is necessary to formulate the S. equation in three dimensions. As for the equation in one dimension, we employ as a starting point the equation for the propagation of a wave motion, which with rectangular axes has the form

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}=\nabla^2\psi=\frac{1}{u^2}\cdot\frac{\partial^2\psi}{\partial t^2}, \tag{1}$$

where ψ is the amplitude, ∇^2 is designated the *Laplacian differential operator*, and u is the velocity of propagation of the wave. This is a partial differential equation of the second order, and the complete integral must define ψ as a function of the three coördinate variables and of the time t.

As in solving equation (2.43), we assume a solution of the form

$$\psi(x, y, z, t) = \phi(x, y, z)\cdot\epsilon^{-2\pi i\nu t},$$

where $\nu\lambda = u$, and consequently

$$\frac{\partial^2\psi}{\partial t^2} = -4\pi^2\nu^2\cdot\phi(x, y, z)\cdot\epsilon^{-2\pi i\nu t}$$

$$= -4\pi^2\nu^2\psi.$$

Hence,

$$\frac{1}{u^2}\cdot\frac{\partial^2\psi}{\partial t^2}=-\frac{4\pi^2\psi}{\lambda^2}=-\frac{4\pi^2}{\lambda^2}\phi(x, y, z)\cdot\epsilon^{-2\pi i\nu t}.$$

Since

$$\frac{\partial^2\psi}{\partial x^2}=\frac{\partial^2\phi}{\partial x^2}\cdot\epsilon^{-2\pi i\nu t}$$

and similar relations apply for $\partial^2\psi/\partial y^2$ and $\partial^2\psi/\partial z^2$, it is evident that the solution of equation (1) satisfies the solution of the partial

differential equation

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}+\frac{4\pi^2}{\lambda^2}\cdot\phi=0, \tag{2a}$$

or

$$\nabla^2\phi+\frac{4\pi^2}{\lambda^2}\cdot\phi=0. \tag{2b}$$

By substituting in the last equation de Broglie's relation for λ, we would obtain the corresponding S. equation in terms of rectangular coördinates. However, in problems involving rotation about an axis of symmetry, or motion of a particle in a central orbit (e.g., the motion of an electron about a positively charged nucleus), it is much more convenient to express the Laplacian operator in *spherical coördinates.* This type of coördinate system is illustrated in Fig. 22*a*, Chapter IV, and the relations between the coördinates r, θ, η and the rectangular coördinates were given in the equations in (*4*.19).

The object of such a transformation of coördinates is, as will be shown by the subsequent argument, to express the S. equation in such a form as will make it possible to *separate the partial differential equation into three ordinary differential equations,* each corresponding to one of the three generalized coördinates. The method is thus analogous to that used in Chapter IV for the solution of the problem of the hydrogen-like atom in terms of spherical coördinates.

6.2 Rules for Transformation of Coördinates. In Chapter IV, section 3, it was mentioned that spherical polar coördinates constitute a special case of a more general class known as *orthogonal curvilinear* systems of coördinates.[1] It has been demonstrated that the three-dimensional S. equation can be separated only in those cases in which it is expressed in terms of such generalized coördinates. Designating these by q_1, q_2, and q_3, the element of distance ds is given, according to equation (*4*.22), by a relation of the form

$$(ds)^2=a_1(dq_1)^2+a_2(dq_2)^2+a_3(dq_3)^2,$$

where each of the coefficients a_1, a_2, and a_3 is in general a function of q_1, q_2, and q_3.

By the application of the theory of vectors[2] or by the application of the methods of the calculus of variations,[3] it is shown that in terms

[1] See Pauling and Wilson, "Introduction to Quantum Mechanics," pp. 104 and 443, for further discussion.

[2] See Appendix IV, Section 13.

[3] Courant and Hilbert, p. 194.

of the generalized coördinates

$$\nabla^2\phi = \frac{1}{\sqrt{a_1a_2a_3}}\left[\frac{\partial}{\partial q_1}\left(\frac{\sqrt{a_1a_2a_3}}{a_1}\cdot\frac{\partial\phi}{\partial q_1}\right)+\frac{\partial}{\partial q_2}\left(\frac{\sqrt{a_1a_2a_3}}{a_2}\cdot\frac{\partial\phi}{\partial q_2}\right)\right.$$
$$\left.+\frac{\partial}{\partial q_3}\left(\frac{\sqrt{a_1a_2a_3}}{a_3}\cdot\frac{\partial\phi}{\partial q_3}\right)\right], \quad (3)$$

and that the element of volume $d\tau$ is given by

$$d\tau = dxdydz = \sqrt{a_1a_2a_3}\cdot dq_1dq_2dq_3,$$

where the expression $\sqrt{a_1a_2a_3}$ is known as the *discriminant.*

It will be observed that since a_1, a_2, and a_3 may each be functions of q_1, q_2, and q_3, the order of operations is important. That is, $\frac{\partial}{\partial q_1}\left(\frac{\sqrt{a_1a_2a_3}}{a_1}\cdot\frac{\partial\phi}{\partial q_1}\right)$ is not identical with $\frac{\sqrt{a_1a_2a_3}}{a_1}\cdot\frac{\partial^2\phi}{\partial q_1^2}$. The *operators are non-commutative,* and in deriving the S. equation this is of extreme significance.

Now in the case of spherical coördinates, as stated in Chapter IV, we have the following relations

$$q_1 = r; \quad q_2 = \theta; \quad q_3 = \eta,$$

and

$$a_1 = 1; \quad a_2 = r^2; \quad a_3 = r^2\sin^2\theta.$$

Hence

$$\sqrt{a_1a_2a_3} = r^2\sin\theta,$$

and therefore

$$d\tau = r^2\sin\theta d\theta d\eta dr.$$

Consequently equation (3) assumes the form

$$\nabla^2\phi = \frac{1}{r^2\sin\theta}\left\{\frac{\partial}{\partial r}\left(r^2\sin\theta\cdot\frac{\partial\phi}{\partial r}\right)+\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial\phi}{\partial\theta}\right)+\frac{\partial}{\partial\eta}\left(\frac{1}{\sin\theta}\cdot\frac{\partial\phi}{\partial\eta}\right)\right\}. \quad (4)$$

It is evident that

$$\frac{\partial}{\partial r}\left(r^2\sin\theta\cdot\frac{\partial\phi}{\partial r}\right) = \sin\theta\cdot\frac{\partial}{\partial r}\left(r^2\cdot\frac{\partial\phi}{\partial r}\right)$$
$$= \sin\theta\left(2r\cdot\frac{\partial\phi}{\partial r}+r^2\cdot\frac{\partial^2\phi}{\partial r^2}\right),$$

while

$$\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial\phi}{\partial\theta}\right) = \cos\theta\cdot\frac{\partial\phi}{\partial\theta}+\sin\theta\cdot\frac{\partial^2\phi}{\partial\theta^2}$$

and $$\frac{\partial}{\partial\eta}\left(\frac{1}{\sin\theta}\cdot\frac{\partial\phi}{\partial\eta}\right)=\frac{1}{\sin\theta}\cdot\frac{\partial^2\phi}{\partial\eta^2},$$

where $\phi = \phi(r, \theta, \eta)$.

Thus, for periodic motions of a particle in three dimensions, equation (2*b*) with the value of the Laplacian operator given in equation (4) takes the place, in quantum mechanics, of the energy equation used in ordinary mechanics, which is of the form

$$E = \tfrac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\cdot\dot{\eta}^2) + V(r, \theta, \eta),$$

where the first term, as indicated in equation (4.37), expresses the kinetic energy, and the second term, the potential energy as a function of the coördinate variables.

This equation is customarily written in the Hamiltonian form

$$E = H(p, q) = \frac{1}{2\mu}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\eta^2}{r^2\sin^2\theta}\right) + V(r, \theta, \eta), \tag{5}$$

and is the form which equation (4.88) assumes for the general case.

Under certain conditions equation (4) assumes simpler forms. Thus, for *constant value of* r (rotation of a sphere about an axis), $\partial/\partial r = 0$, and the Laplacian term becomes

$$\nabla^2\phi = \frac{1}{r^2\sin\theta}\cdot\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial\phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\cdot\frac{\partial^2\phi}{\partial\eta^2}. \tag{6}$$

For the case of plane polar coördinates (r, η) the Laplacian may be derived from the relations

$$x = r\cos\eta$$
$$y = r\sin\eta$$

as follows:[4]

Since $r = \sqrt{x^2 + y^2}$, and $\tan\eta = y/x$, therefore,

$$\frac{\partial r}{\partial x} = \frac{x}{r}; \quad \frac{\partial r}{\partial y} = \frac{y}{r}; \quad \frac{\partial\eta}{\partial x} = -\frac{y}{r^2}; \quad \frac{\partial\eta}{\partial y} = \frac{x}{r^2};$$

$$\frac{\partial^2 r}{\partial x^2} = \frac{1}{r} - \frac{x}{r^2}\cdot\frac{\partial r}{\partial x} = \frac{1}{r} - \frac{x^2}{r^3};$$

$$\frac{\partial^2 r}{\partial y^2} = \frac{1}{r} - \frac{y^2}{r^3}; \quad \frac{\partial^2\eta}{\partial x^2} = \frac{\partial}{\partial x}\left(-\frac{y}{r^2}\right) = \frac{2xy}{r^4};$$

$$\frac{\partial^2\eta}{\partial y^2} = -\frac{2xy}{r^4}.$$

[4] Slater and Frank, "Introduction to Theoretical Physics," pp. 164–165.

Now,

$$\frac{\partial\phi}{\partial x}=\frac{\partial\phi}{\partial r}\cdot\frac{\partial r}{\partial x}+\frac{\partial\phi}{\partial\eta}\cdot\frac{\partial\eta}{\partial x}.$$

Consequently,

$$\frac{\partial^2\phi}{\partial x^2}=\frac{\partial^2\phi}{\partial r^2}\left(\frac{\partial r}{\partial x}\right)^2+2\frac{\partial^2\phi}{\partial r\partial\eta}\left(\frac{\partial r}{\partial x}\cdot\frac{\partial\eta}{\partial x}\right)+\frac{\partial^2\phi}{\partial\eta^2}\left(\frac{\partial\eta}{\partial x}\right)^2$$

$$+\frac{\partial\phi}{\partial r}\cdot\frac{\partial^2 r}{\partial x^2}+\frac{\partial\phi}{\partial\eta}\cdot\frac{\partial^2\eta}{\partial x^2},$$

and a similar equation will be obtained for $\frac{\partial^2\phi}{\partial y^2}$. Adding these two equations, we have

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}=\frac{\partial^2\phi}{\partial r^2}\left[\left(\frac{\partial r}{\partial x}\right)^2+\left(\frac{\partial r}{\partial y}\right)^2\right]+2\frac{\partial^2\phi}{\partial r\partial\theta}\left(\frac{\partial r}{\partial x}\cdot\frac{\partial\eta}{\partial x}+\frac{\partial r}{\partial y}\cdot\frac{\partial\eta}{\partial y}\right)$$

$$+\frac{\partial^2\phi}{\partial\eta^2}\left[\left(\frac{\partial\eta}{\partial x}\right)^2+\left(\frac{\partial\eta}{\partial y}\right)^2\right]+\frac{\partial\phi}{\partial r}\left(\frac{\partial^2 r}{\partial x^2}+\frac{\partial^2 r}{\partial y^2}\right)$$

$$+\frac{\partial\phi}{\partial\theta}\left(\frac{\partial^2\eta}{\partial x^2}+\frac{\partial^2\eta}{\partial y^2}\right).$$

Substituting from the first set of equations, the result is

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\phi=\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r^2}\cdot\frac{\partial^2}{\partial\eta^2}+\frac{1}{r}\cdot\frac{\partial}{\partial r}\right)\phi.$$

It will be observed that this result could also have been obtained directly from the rule of transformation in equation (3) thus:

Since $(ds)^2=(dr)^2+r^2(\partial\eta)^2$,

$$a_1=1,\ a_2=r^2;\quad \sqrt{a_1a_2}=r.$$

Therefore,

$$\nabla^2\phi=\frac{1}{r}\left\{\frac{\partial}{\partial r}\left(\frac{r}{1}\cdot\frac{\partial\phi}{\partial r}\right)+\frac{\partial}{\partial\eta}\left(\frac{r}{r^2}\cdot\frac{\partial\phi}{\partial\eta}\right)\right\}$$

$$=\frac{\partial^2\phi}{\partial r^2}+\frac{1}{r}\cdot\frac{\partial\phi}{\partial r}+\frac{1}{r^2}\cdot\frac{\partial^2\phi}{\partial\eta^2}. \qquad (7a)$$

In the case of radial motion, for which $\partial/\partial\theta = \partial/\partial\eta = 0$, the Laplacian operator evidently assumes the form

$$\nabla^2 = \frac{1}{r^2} \cdot \frac{d}{dr}\left(r^2 \frac{d}{dr}\right) = \frac{d^2}{dr^2} + \frac{1}{r} \cdot \frac{d}{dr} \tag{7b}$$

We shall now return to the consideration of equation (2*b*). If we express $\nabla^2\phi$ in this equation in terms of spherical polar coördinates and substitute for λ, the de Broglie relation

$$\lambda = \frac{h}{\sqrt{2\mu(E - U)}},$$

we obtain the S. equation of the form

$$\nabla^2\phi + \frac{8\pi^2\mu}{h^2}(E - U)\phi = 0, \tag{8}$$

where U, the potential energy, is a function of r, θ, and η. As in the case of the S. equation for one coördinate variable, we seek solutions of equation (8) that will be physically rational. Thus $\phi\bar{\phi}$ must not become infinite at any point in space, and it must tend to vanish as those regions are approached in which the probability of occurrence tends to become zero. The exact form of these "boundary conditions" must depend upon the nature of the particular problem. Thus, in the case of a hydrogen-like atom, the probability of occurrence of the electron must decrease continuously to zero as r tends toward infinitely large values. We shall find that actually this probability becomes infinitesimally small for values of r exceeding only a few atomic radii. In the case of the angle variables, the limits are $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \eta \leqslant 2\pi$, and the distribution function, as we shall designate $\phi\bar{\phi}$, must exhibit a periodicity with respect to these variables. That is, $\phi\bar{\phi}(\theta, \eta) = \phi\bar{\phi}(\theta \pm \pi, \eta \pm 2\pi)$.

Furthermore, because experimental observations show that any atomic or molecular system can exist only in a series of discrete states defined by the energy values E_1, E_2 . . . and so forth, we must expect, if the solutions of the S. equation correspond to the observations, that "sensible" solutions of equation (8) will exist only for a series of discrete values of the energy E, which will constitute the *eigenvalues*. The corresponding eigenfunctions ϕ will represent, in the most general case, amplitudes of stationary de Broglie waves in three dimensions, and cannot therefore be visualized physically. In the case of constant values of r, ϕ represents the amplitude of vibration of a spherical surface, and hence the functions are known as *surface spherical harmonics*. They are represented by expressions which are functions of the latitude

(θ) and longitude (η) and which exhibit *nodes* and *loops* along both meridian circles and zonal circles (parallel to the equatorial plane). Consequently, the mathematical expressions are quite complicated and, in fact, appear formidable at first glance.

To some, indeed, it might appear that the mathematician has endowed nature with a complexity far beyond its needs. Yet the only reply to such an accusation must be that the " simple " solutions, those which are relatively easy to understand (because they involve no " higher mathematics "), do not correspond to the facts. Nature is complex in its fundamental elements, and the only feature that is astounding is this: that human intelligence has been able to devise a method of reasoning with symbols by which a one-to-one correspondence is attained between the deductions from this reasoning and the experimental facts. This, to the mind of the writer, has always appeared the most marvelous aspect of all mathematical technic in dealing with the interpretation of nature. And because it is stimulating to understand this " picture "; because it must add a certain measure of pleasure to perceive, even though it be dimly, at first, the results attained by combining transcendental imagination with the most exacting type of logic — because of these rewards which the effort holds forth — let the reader not be discouraged too readily. Patience and persistence alone will accomplish wonders, even in the comprehension of a symbolic mathematical technic.

6.3 The Rigid Rotator with Fixed Axis. Let us consider the problem of a diatomic molecule constituted of two atoms of masses μ_1 and μ_2, located at distances r_1 and r_2 from the axis of rotation. We shall assume that the molecule is of the " dumbbell " form, so that the distance between the centers of the atoms ($r_1 + r_2 = r_0$) is fixed, and therefore we neglect the possibility that the atoms will vibrate along this axis in virtue of their mutual attractive and repulsive forces. (Of course, such vibrations, of frequency $n\nu_0$, actually occur and give rise to vibrational energy states — a problem which was considered in the case of the harmonic oscillator.) Under these conditions we may regard the molecule as possessing, in general, *two degrees of freedom* or mobility. The molecule will have a rotational motion about an axis of symmetry passing through its center of gravity. This will be represented by an angular velocity $\dot{\eta} = d\eta/dt$ in the plane YOX (see Fig. 22*a*). Also there will be a precessional motion of the fixed axis of the molecule about the axis of symmetry, which is represented by the angular velocity $\dot{\theta}$. Since there is no potential energy term, the total energy is all kinetic and is given by

$$E = \tfrac{1}{2}(\mu_1 r_1^2 + \mu_2 r_2^2)\,(\dot{\theta}^2 + \sin^2\theta \cdot \dot{\eta}^2). \tag{9}$$

Since the molecule is rotating about its center of gravity,

$$\mu_1 r_1 = \mu_2 r_2.$$

Hence, if we put

$$r_1 = \frac{\mu_2 r_0}{\mu_1 + \mu_2}; \quad r_2 = \frac{\mu_1 r_0}{\mu_1 + \mu_2},$$

then

$$\mu_1 r_1^2 + \mu_2 r_2^2 = \frac{\mu_1 \mu_2}{\mu_1 + \mu_2} \cdot r_0^2 = \mu r_0^2 = I,$$

and we can write (9) in the form

$$E = \frac{I}{2}(\dot{\theta}^2 + \sin^2\theta \cdot \dot{\eta}^2), \tag{10}$$

where I = moment of inertia of the molecule about its center of gravity, r_0 = mean radius of gyration, and $1/\mu = 1/\mu_1 + 1/\mu_2$, where μ is known as the " reduced " mass.

Thus $r_0^2(\dot{\theta}^2 + \sin^2\theta \cdot \dot{\eta}^2) = v^2$, where v is the velocity of rotation, and the corresponding de Broglie wave length is given by

$$\lambda = \frac{h}{\mu v} = \frac{h}{\sqrt{2\mu E}} = \frac{h r_0}{\sqrt{2EI}}.$$

Hence, the S. equation

$$\nabla^2\phi + \frac{4\pi^2\phi}{\lambda^2} = 0$$

becomes

$$\nabla^2\phi + \frac{8\pi^2 EI\phi}{h^2 r_0^2} = 0.$$

Since I is a constant for any given diatomic molecule in the state for which the energy is E, it follows that both v and λ are constants.

Furthermore, it follows that we may use the form of the Laplacian operator given in (6) with $r^2 = r_0^2$. Hence, multiplying both terms by r_0^2 the equation to be solved is

$$\frac{1}{\sin\theta} \cdot \frac{\partial}{\partial\theta}\left(\sin\theta \cdot \frac{\partial\phi}{\partial\theta}\right) + \frac{1}{\sin^2\theta} \cdot \frac{\partial^2\phi}{\partial\eta^2} + \frac{8\pi^2 EI\phi}{h^2} = 0. \tag{11}$$

There are two cases which may occur. In the first of these, the molecule is free to revolve only about an axis at right angles to the axis of the molecule. This is known as the case of the rigid rotator with fixed axis.

In the second case, the molecule may exhibit a motion of precession, as well as that of rotation. The latter is the case for which equation (11) applies and will be discussed in a following section. In the first case, however, $\partial/\partial\theta = 0$, $\sin\theta = 1$, and the equation reduces to the ordinary differential equation

$$\frac{d^2\phi}{d\eta^2} + m^2\phi = 0, \tag{12}$$

where $m^2 = 8\pi^2 EI/h^2$ is used to indicate that the coefficient is always positive.

The solution of this equation has been discussed previously. It is

$$\phi = A\epsilon^{im\eta} + B\epsilon^{-im\eta}, \tag{13}$$

which may be written in the form of a sine or cosine function. Thus we may use the form

$$\phi = C\sin(m\eta + \delta),$$

where δ is a phase angle.

Now this equation has physical significance only if $m = 0, 1, 2$, etc. Consequently, the *eigenvalues* of the discrete energy states are given by the relation

$$E_m = \frac{m^2h^2}{8\pi^2 I}, \tag{14}$$

where m is an arbitrary integer.

Now from equations (10) and (14) it follows, for the case $\partial/\partial\theta = 0$, that

$$2E_m I = \frac{m^2h^2}{4\pi^2}$$

$$= I^2\dot{\eta}^2.$$

Hence,

$$I\dot{\eta} = \pm\frac{mh}{2\pi}.$$

That is, the angular momentum of rotation of the molecule is equal to an integral multiple of $h/2\pi$, and the plus and minus signs refer to opposite directions of rotation. Thus, the interpretation of $\epsilon^{\pm im\eta}$ is analogous to that of $\epsilon^{\pm i\alpha x}$ in the case of wave propagation along the x-axis. (See Chapter II.)

If we set

$$A\bar{A}\int_0^{2\pi}\epsilon^{im\eta}\epsilon^{-im\eta}d\eta = 1,$$

it is evident that

$$A\bar{A} = \frac{1}{2\pi}.$$

Consequently, corresponding to any one eigenvalue E_m, we have the *two* normalized eigenfunctions,

$$\phi_m = \frac{1}{\sqrt{2\pi}} \epsilon^{im\eta}, \tag{15a}$$

and

$$\phi_{-m} = \frac{1}{\sqrt{2\pi}} \epsilon^{-im\eta}. \tag{15b}$$

Obviously, the functions ϕ_n and $\bar{\phi}_m$, $n \neq m$, are orthogonal, since, as shown in Chapter III,

$$\int_0^{2\pi} \epsilon^{i(m-n)\eta} d\eta = \int_0^{2\pi} \cos (m - n)\eta d\eta + i \int_0^{2\pi} \sin (m - n)\eta d\eta = 0.$$

It will be observed that in this case there are *two eigenfunctions*, given by equations (15*a*) and (15*b*), corresponding to any given eigenvalue. We have here an illustration of a condition that is met with frequently in the solution of problems in quantum mechanics. Such energy states, for which there are available more than one eigenfunction for any given eigenvalue, are known as *degenerate* (German, " *entartet* "). Physically this is interpreted, in the present case, as indicating that actually there are two energy states which have become merged (degenerated) into what appears to be one state, because the energy is the same *irrespective of the direction of rotation of any molecule with respect to other molecules.* However, if the molecules are placed in a magnetic field, the energy will vary (because the molecules possess magnetic moments) with the direction of rotation of the molecule. For one direction of rotation the energy will be slightly greater, and for the opposite direction slightly less, than the value E_m which exists in the absence of a magnetic field.

6.4 The Rigid Rotator with Free Axis. We shall now consider the case of the *rigid rotator with free axis*, for which the S. equation is that given in (11). The solution must represent ϕ as $\phi(\theta, \eta)$, that is, as a function of the two angle variables. To solve the partial differential equation, we postulate a solution for ϕ of the form

$$\phi(\theta, \eta) = X(\theta) \cdot Z(\eta),$$

where X is a function of θ only, and Z of η only. We have an indication that this is possible from the fact that in the case of the rotator with fixed axis we have already found a limiting case of this problem. Evidently,

$$\frac{\partial^2 \phi}{\partial \theta^2} = Z(\eta) \cdot \frac{d^2 X(\theta)}{d\theta^2},$$

and

$$\frac{\partial^2\phi}{\partial\eta^2} = X(\theta) \cdot \frac{d^2Z(\eta)}{d\eta^2}.$$

Substituting for ϕ and its derivatives in equation (11) and using the symbol

$$\alpha^2 = \frac{8\pi^2 EI}{h^2}, \tag{16}$$

we obtain the relation

$$\frac{Z}{\sin\theta} \cdot \frac{d}{d\theta}\left(\sin\theta \cdot \frac{dX}{d\theta}\right) + \frac{X}{\sin^2\theta} \cdot \frac{d^2Z}{d\eta^2} + \alpha^2 XZ = 0,$$

where $Z = Z(\eta)$ and $X = X(\theta)$.

Since $\sin^2\theta/(XZ)$ never becomes infinite, we can multiply through by this factor and thus obtain the equation

$$\frac{\sin\theta}{X} \cdot \frac{d}{d\theta}\left(\sin\theta \cdot \frac{dX}{d\theta}\right) + \alpha^2 \sin^2\theta = -\frac{1}{Z} \cdot \frac{d^2Z}{d\eta^2}.$$

It will be observed that the left-hand side does not involve η, and the right-hand side does not involve θ. Since this relation must be valid for all possible values of θ and η, it follows that *each side of the equation must be equal to a constant*, which we shall designate by m^2. We thus obtain the two ordinary differential equations

$$\frac{d^2Z}{d\eta^2} + m^2Z = 0 \tag{17}$$

and

$$\sin\theta \cdot \frac{d}{d\theta}\left(\sin\theta \cdot \frac{dX}{d\theta}\right) + (\alpha^2 \sin^2\theta - m^2)X = 0. \tag{18}$$

The first of these is identical with equation (12). Therefore the solutions are given by

$$Z_m = \frac{\epsilon^{\pm im\eta}}{\sqrt{2\pi}}, \tag{19}$$

where $m = 0, 1, 2$, etc.

Now let us consider equation (18), and as a first step in the process of solving it we change to a variable x, such that

$$x = \cos\theta.$$

Therefore,

$$1 - x^2 = \sin^2\theta,$$

and

$$\frac{d}{d\theta} = -\sin\theta \cdot \frac{d}{dx}.$$

Before carrying through the transformation to the new variable, we may divide through by $\sin^2\theta$. This gives

$$\frac{1}{\sin\theta} \cdot \frac{d}{d\theta}\left(\sin\theta \cdot \frac{dX}{d\theta}\right) + \left(\alpha^2 - \frac{m^2}{\sin^2\theta}\right)X = 0.$$

Introducing the variable x, this becomes

$$\frac{d}{dx}\left\{(1 - x^2)\frac{dX}{dx}\right\} + \left(\alpha^2 - \frac{m^2}{1 - x^2}\right)X = 0. \tag{20}$$

This equation is one of the most important in mathematical physics and is known as Legendre's equation of *order* m, in x, where $-1 \leqslant x \leqslant 1$. That is, the equation has physical significance between the limits $x = \cos\theta = \pm 1$. These limits constitute so-called singular points, since $1 - x^2 = 0$ at these points.

Since m can have any integral value, including 0, we shall consider first the solution of equation (20) for the case $m = 0$, that is, the *Legendre equation of order zero*,

$$\frac{d}{dx}\left\{(1 - x^2)\frac{dX}{dx}\right\} + \alpha^2 X = 0,$$

or

$$(1 - x^2)\frac{d^2X}{dx^2} - 2x \cdot \frac{dX}{dx} + \alpha^2 X = 0. \tag{21}$$

Let us assume, as in equation (5.13) for the linear oscillator, that X may be represented as a polynomial of degree k, so that

$$X = \sum a_k x^k, \tag{22}$$

where

$$k = 0, 1, 2, \ldots (k - 1), k.$$

Then

$$\frac{dX}{dx} = \sum a_k \cdot k \cdot x^{k-1}$$

$$-2x \cdot \frac{dX}{dx} = -\sum 2a_k \cdot k \cdot x^k.$$

Also, in terms of coefficients of x^k,

$$\frac{d^2X}{dx^2} = \sum a_{k+2}(k+2)\,(k+1)\cdot x^k,$$

and

$$-x^2\frac{d^2X}{dx^2} = -\sum a_k \cdot k(k-1)x^k.$$

Therefore, the coefficient of x^k in (21) is given by

$$a_{k+2}(k+2)\,(k+1) - a_k\{k(k-1) + 2k - \alpha^2\}.$$

Since equation (21) is valid for all values of x (in the range $-1 \leqslant x \leqslant 1$), it follows that the coefficient of each power of x must vanish identically. Hence,

$$a_{k+2} = \frac{a_k\{k(k+1) - \alpha^2\}}{(k+2)\,(k+1)}. \tag{23}$$

Thus

$$\frac{a_{k+2}}{a_k} = \frac{k}{k+2} - \frac{\alpha^2}{(k+1)\,(k+2)}.$$

If k can increase beyond limit, $a_{k+2}/a_k = 1$ for very large values of k, and consequently, if the series for X defined by equation (22) is to converge for $x = \pm 1$, it must have a finite number of terms and the highest power of x is given by k, where $a_{k+2} = 0$. Consequently,

$$k(k+1) = \alpha^2. \tag{24}$$

Substituting for α^2 from equation (16) it follows that E can assume only the series of discrete values given by the relation

$$E_k = \frac{k(k+1)h^2}{8\pi^2 I}, \tag{25}$$

where $k = 0, 1, 2,$ etc.

This relation is different from that deduced for the rotator with rigid axis which was stated in equation (14), and is actually in much better agreement with the spectroscopic observations on the rotational energy levels of molecules than the latter, which is identical with the relation derived by means of classical mechanics.

6.5 Legendre Equation of Order Zero. Substituting in equation (21) the value for α^2 deduced in equation (24), we obtain the differential equation for X_k, the eigenfunction corresponding to E_k, in the form

$$(1 - x^2)X_k'' - 2x \cdot X_k' + k(k+1)X_k = 0.$$

From (23) it follows that

$$a_k = \frac{\{(k-2)(k-1) - k(k+1)\}}{k(k-1)} \cdot a_{k-2}$$

$$= -\frac{2(2k-1)}{k(k-1)} \cdot a_{k-2},$$

or

$$a_{k-2} = -\frac{k(k-1)}{2(2k-1)} \cdot a_k. \tag{26}$$

Similarly, it is readily shown that

$$a_{k-4} = -\frac{(k-2)(k-3)}{4(2k-3)} \cdot a_{k-2}$$

$$= \frac{k(k-1)(k-2)(k-3)}{2 \cdot 4 \cdot (2k-1)(2k-3)} \cdot a_k, \tag{27}$$

and

$$a_{k-6} = \frac{k(k-1)(k-2)(k-3)(k-4)(k-5)}{2 \cdot 4 \cdot 6 \cdot (2k-1)(2k-3)(2k-5)} \cdot a_k. \tag{28}$$

If k is even, the power series beginning with x^k will end with a_0, if k is odd, with $a_1 x$. We thus obtain the series,

$$X_k = a_k \left[x^k + \frac{a_{k-2} x^{k-2}}{a_k} + \frac{a_{k-4} x^{k-4}}{a_k} + \ldots \right], \tag{29}$$

where the coefficients of the various powers of x are given by relations similar to (28), and a_k is arbitrary.

If we assign the value[5]

$$a_k = \frac{(2k-1)(2k-3)\ldots 1}{k!}$$

the resulting function X_k is known as a *Legendre function of order zero and degree* k or a *surface zonal harmonic* and designated by the symbol $P_k(x) = P_k(\cos\theta)$. Thus the complete expression for the Legendre function has the form

$$P_k(x) = \frac{(2k-1)(2k-3)\ldots 1}{k!} \left[x^k - \frac{k(k-1)}{2(2k-1)} \cdot x^{k-2} \right.$$

$$\left. + \frac{k(k-1)(k-2)(k-3)}{2 \cdot 4 \cdot (2k-1)(2k-3)} \cdot x^{k-4} - \ldots \right]. \tag{30}$$

[5] This value, as will appear from (31), makes the first Legendre function $P_0(x) = 1$.

The first few members of this series are as follows:[6]

$$
\begin{aligned}
P_0(x) &= 1 && \text{or } P_0(\cos\theta) = 1\\
P_1(x) &= x && \text{or } P_1(\cos\theta) = \cos\theta\\
P_2(x) &= \tfrac{1}{2}(3x^2 - 1) && \text{or } P_2(\cos\theta) = \tfrac{1}{2}(3\cos^2\theta - 1)\\
P_3(x) &= \tfrac{1}{2}(5x^3 - 3x) && \text{or } P_3(\cos\theta) = \tfrac{1}{2}(5\cos^3\theta - 3\cos\theta)\\
P_4(x) &= \tfrac{1}{8}(35x^4 - 30x^2 + 3)\\
P_5(x) &= \tfrac{1}{8}(63x^5 - 70x^3 + 15x)
\end{aligned}
\qquad (31)
$$

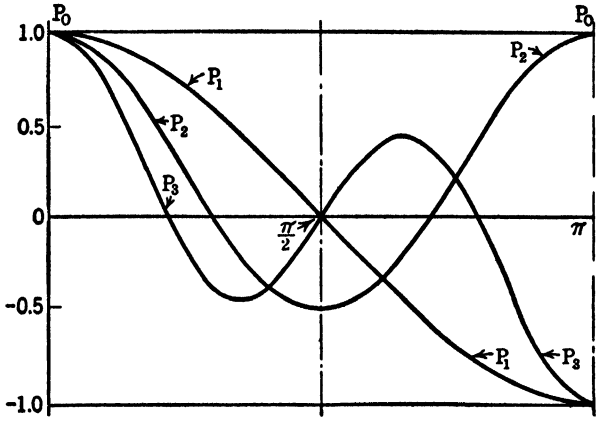


FIG. 26. Plots of the first four Legendre polynomials as functions of θ.

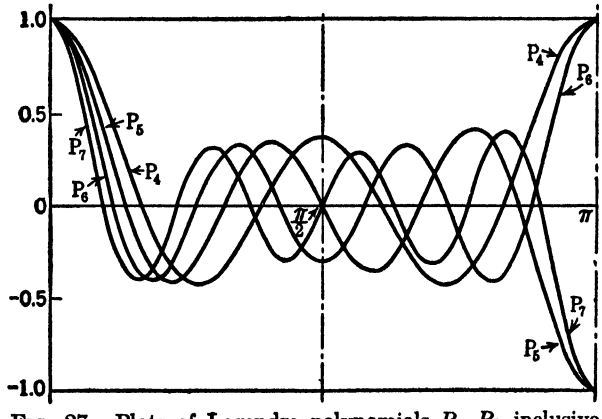


FIG. 27. Plots of Legendre polynomials P_4–P_7 inclusive.

At $\theta = 0$, $\cos\theta = x = 1$, while at $\theta = \pi$, $\cos\theta = -1$, and at $\theta = \pi/2$, $\cos\theta = x = 0$. The functions may be plotted as functions of either θ or x. In the former case, the limits are 0 and π; in the latter, the corresponding limits are 1 and -1. Figures 26 and 27 (taken from the curves plotted in Byerly's treatise) show graphs of the functions $P_1(\cos\theta)$

[6] Tables of values of P_1 to P_7 (inclusive), as functions of θ, are given in L. Silberstein's "Mathematical Tables" and in an appendix in Byerly's "Fourier's series."

to $P_7(\cos\theta)$, inclusive, as functions of θ. It will be observed that at the limits all the functions have the identical absolute value 1. Furthermore, since

$$P_k(-x) = (-1)^k P_k(x),$$

it follows that for k even, $P_k(-x) = P_k(x)$, that is, the function is symmetrical about $x = 0$ or $\theta = \pi/2$, resembling in this respect $\sin\theta$, while for k odd, $P_k(-x) = -P_k(x)$, and the functions change sign in passing through $x = 0$ or $\theta = \pi/2$, which also is characteristic of $\cos\theta$.

As pointed out in connection with the Hermitian polynomials, the number of roots, corresponding to *nodes*, along the axis of x or θ, is equal to that of the highest power of x. Thus $P_1(x) = \cos\theta$ passes through 0 at $\theta = \pi/2$, while $P_2(x)$ exhibits two nodes which may be determined from the quadratic relation

$$P_2(x) = \tfrac{1}{2}(x\sqrt{3} + 1)\,(x\sqrt{3} - 1) = 0.$$

Hence $P_2(x) = 0$ for $\cos\theta = \pm 1/\sqrt{3} = \pm 0.5775$, that is, for $\theta_1 = 54^\circ\,45'$ and $\theta_2 = 125^\circ\,15''$. In a similar manner it is possible to calculate the k values of θ at which any given function $P_k(\cos\theta)$ becomes equal to zero.

As for the other polynomials, it is readily shown in any of the treatises on this topic that the Legendre polynomials are related by a recursion formula of the form

$$(k+1)P_{k+1}(x) = (2k+1)xP_k(x) - kP_{k-1}(x),$$

which makes it possible to calculate higher members of the series from the lower members.

From equation (30) it follows that

$$\int_0^x P_k(x)dx = \frac{(2k-1)\,(2k-3)\ldots 1}{(k+1)!}\left[x^{k+1} - \frac{(k+1)k}{2(2k-1)}\cdot x^{k-1} + \frac{(k+1)k(k-1)\,(k-2)}{2\cdot 4\cdot(2k-1)\,(2k-3)}\cdot x^{k-3} - \ldots\right],$$

and if this integration is repeated $k-1$ times more, the result is

$$\frac{(2k-1)\,(2k-3)\ldots 1}{(2k)!}\left[x^{2k} - \frac{2k(2k-1)}{2\cdot(2k-1)}\cdot x^{2k-2} + \frac{2k(2k-1)\,(2k-2)\,(2k-3)}{2\cdot 4\cdot(2k-1)\,(2k-3)}\cdot x^{2k-4} - \ldots\right]$$

$$= \frac{1}{(2k)\,(2k-2)\ldots 2}(x^2-1)^k = \frac{1}{2^k k!}(x^2-1)^k, \qquad (32)$$

as may be demonstrated by expanding the expression $(x^2 - 1)^k$. The kth derivative of the function on the right-hand side of the last equation is therefore $P_k(x)$, so that

$$P_k(x) = \frac{1}{2^k \cdot k!} \cdot \frac{d^k(x^2 - 1)^k}{dx^k} \cdot \tag{33}$$

This is known as *Rodrigues' Formula.*

It may be shown by means of this formula that the *Legendre polynomials form an orthogonal system,* since

$$\int_{-1}^{1} P_k(x)P_n(x)dx = 0 \text{ (for } k \neq n)$$

$$= \frac{2}{2k + 1} \text{ (for } k = n). \tag{34}$$

Hence, the normalizing factor is

$$N_k = \sqrt{\frac{2}{2k + 1}} \cdot \tag{35}$$

It follows from these relations that if $n + k$ is even and $n \neq k$,

$$\int_0^1 P_k(x)P_n(x)dx = \int_{-1}^0 P_k(x)P_n(x)dx = 0,$$

and also that

$$\int_0^1 \{P_k(x)\}^2 dx = \frac{1}{2k + 1} \cdot$$

The orthogonality relation (34) may also be derived by the following proof which is independent of Rodrigues' formula, and is of very general application in spherical harmonics.

Since $P_k(x)$ and $P_n(x)$ satisfy Legendre's equation,

$$\frac{d}{dx}\left\{(1 - x^2)\frac{dP_k}{dx}\right\} + k(k + 1)P_k = 0, \tag{i}$$

and

$$\frac{d}{dx}\left\{(1 - x^2)\frac{dP_n}{dx}\right\} + n(n + 1)P_n = 0. \tag{ii}$$

Multiplying (i) by P_n and (ii) by P_k, subtracting, and then integrating, we obtain the relation

$$[n(n + 1) - k(k + 1)] \int_{-1}^{1} P_n P_k dx =$$

$$\int_{-1}^{1} P_n \frac{d}{dx}\left[(1 - x^2)\frac{dP_k}{dx}\right] dx - \int_{-1}^{1} P_k \frac{d}{dx}\left[(1 - x^2)\frac{dP_n}{dx}\right] dx.$$

Integrating the right-hand side by parts, it becomes

$$\left[P_n(1-x^2)\cdot\frac{dP_k}{dx}-P_k(1-x^2)\frac{dP_n}{dx}\right]_{x=-1}^{x=1}$$

$$-\int_{-1}^{1}(1-x^2)\frac{dP_k}{dx}\cdot\frac{dP_n}{dx}\cdot dx+\int_{-1}^{1}(1-x^2)\frac{dP_n}{dx}\cdot\frac{dP_k}{dx}\cdot dx,$$

which is equal to zero, since $1-x^2=0$ at $x=\pm 1$, while the two integrals cancel. Consequently, if $n\neq k$,

$$\int_{-1}^{1}P_n(x)P_k(x)dx=0.$$

6.6 Associated Legendre Functions. We shall now consider equation (20) for the case $m\geqq 1$. In the preceding discussion this equation was solved for the case $m=0$, and it was found that the equation gives solutions which are physically significant if we put

$$\alpha^2=k(k+1). \tag{24}$$

If we substitute for α^2 in equation (20) we obtain the relation

$$\frac{d}{dx}\left\{(1-x^2)\frac{dX}{dx}\right\}+\left\{k(k+1)-\frac{m^2}{1-x^2}\right\}X=0. \tag{36}$$

Let us introduce a new function Y defined by the relation

$$X=(1-x^2)^{\frac{m}{2}}Y. \tag{37}$$

Then

$$\frac{dX}{dx}=\frac{-mx(1-x^2)^{\frac{m}{2}}Y}{(1-x^2)}+\frac{(1-x^2)^{\frac{m}{2}+1}}{(1-x^2)}\cdot\frac{dY}{dx};$$

$$(1-x^2)\frac{dX}{dx}=-mx(1-x^2)^{\frac{m}{2}}Y+(1-x^2)^{\frac{m}{2}+1}\cdot\frac{dY}{dx},$$

and

$$\frac{d}{dx}\left\{(1-x^2)\frac{dX}{dx}\right\}=\left\{-m(1-x^2)^{\frac{m}{2}}+\frac{m^2x^2(1-x^2)^{\frac{m}{2}}}{(1-x^2)}\right\}Y$$

$$-\left\{mx(1-x^2)^{\frac{m}{2}}+2x\left(\frac{m}{2}+1\right)(1-x^2)^{\frac{m}{2}}\right\}\frac{dY}{dx}+(1-x^2)^{\frac{m}{2}}+\frac{d^2Y}{dx^2}\cdot$$

Substituting these relations into (36) and dividing by $(1-x^2)^{\frac{m}{2}}$,

which is not zero, except at the limits, we obtain the relation

$$(1 - x^2)\frac{d^2Y}{dx^2} - 2(m + 1)x \cdot \frac{dY}{dx} + (k - m)(k + m + 1)Y = 0. \quad (38)$$

With this differential equation we shall compare the differential equation (21) which can evidently be written in the form

$$(1 - x^2)\frac{d^2P_k}{dx^2} - 2x \cdot \frac{dP_k}{dx} + k(k + 1)P_k = 0.$$

If we now differentiate this m times, it can be shown that the resulting equation is identical with (38), and consequently,

$$Y = \frac{X}{(1 - x^2)^{\frac{m}{2}}} = \frac{d^mP_k}{dx^m}. \quad (39)$$

The function X thus obtained, see equation (37), is known as the *associated Legendre Function of the first kind, of degree* k *and order* m, and is designated thus

$$P_k^m(x) = (1 - x^2)^{\frac{m}{2}}\frac{d^mP_k}{dx^m}. \quad (40)$$

Since the differential coefficient becomes zero for $m > k$, it follows that m can have only the series of integral values $m = 0, 1, 2, \ldots k$. Thus corresponding to any value of k, there are $(k + 1)$ Legendre functions which satisfy equation (20) and also $(2k + 1)$ functions Z_m, which satisfy equation (17), corresponding to $m = 0, \pm 1, \pm 2, \ldots \pm k$.

Since

$$(1 - x^2)^{\frac{m}{2}} = \sin^m \theta,$$

the associated Legendre function may also be written in the form[7]

$$P_k^m(\cos\theta) = \sin^m\theta\frac{d^mP_k(\cos\theta)}{d(\cos\theta)^m}. \quad (41)$$

[7] The notation given is that used by Byerly, MacRobert, Pauling and Wilson, and most of the authorities. On the other hand, Courant and Hilbert, as well as Condon and Morse, use the notation

$$P_k^m(x) = \frac{d^mP_k(x)}{dx^m}$$

and

$$P_{km}(x) = \sin^m\theta \cdot P^m(x).$$

The functions

$$\left.\begin{aligned} &\cos m\eta \sin^m \theta \cdot \frac{d^m P_k(x)}{dx^m} \\ \text{and} \\ &\sin m\eta \sin^m \theta \cdot \frac{d^m P_k(x)}{dx^m} \end{aligned}\right\} \tag{42}$$

are known as tesseral harmonics of the k*th degree and* m*th order.* In terms of exponentials, the functions are

$$\epsilon^{im\eta} \sin^m \theta \cdot \frac{d^m P_k(x)}{dx^m} \quad \text{and} \quad \epsilon^{-im\eta} \sin^m \theta \cdot \frac{d^m P_k(x)}{dx^m}, \tag{43}$$

and as stated already, for any given value of k, there are $(2k + 1)$ *functions which satisfy the differential equations* (17) *and* (18). It will be observed that the only condition attached to m in the solution of (17) is that it must have an integral value (including 0). The condition that m cannot exceed k was derived from the subsequent deduction based on the fact that the eigenfunctions which satisfy equation (18) are of the type $P_k^m(x)$.

It was also deduced that the eigenvalues corresponding to the different energy states are given by

$$E_k = \frac{k(k+1)h^2}{8\pi^2 I}. \tag{25}$$

Thus it follows that for *any given energy state, corresponding to a given value of* k, *there are* (2k + 1) *eigenfunctions.* In the case of the rotator with fixed axis, it was found that for each energy state there are two possible eigenfunctions. The order of degeneracy is therefore two in that case. But in *the case of the rotator with free axis we find that the order of degeneracy is* 2k + 1. The physical interpretation is similar to that given in the previous case. As Condon and Morse describe it: " this is the degeneracy of random space orientation in a centrally symmetric field, and gives the multiplicity into which the terms are split when a non-symmetric perturbing field removes the degeneracy." Thus in the presence of magnetic or electrostatic fields this degeneracy may be completely removed, because such fields introduce perturbing effects.

The functions $P_k^m(x)$ satisfy the condition for orthogonality of the form

$$\int_{-1}^{1} P_k^m(x) \cdot P_j^m(x) dx = 0. \quad (j \neq k), \tag{44}$$

where obviously m must not exceed either j or k. This may be deduced by an argument quite similar to that used for demonstrating the orthogonal nature of the Legendre polynomials of zero order. That is any two Legendre polynomials of the same order and different degrees are orthogonal to each other.

It may also be shown that

$$\int_{-1}^{1} \{P_k^m(x)\}^2 dx = \frac{2}{2k+1} \cdot \frac{(k+m)!}{(k-m)!}. \tag{45}$$

Hence the normalizing factor for the tesseral harmonic given in (42) or (43) is given by[8]

$$N = \left\{ \frac{2\pi}{1} \cdot \frac{2}{(2k+1)} \cdot \frac{(k+m)!}{(k-m)!} \right\}^{\frac{1}{2}}. \tag{46}$$

Finally, it is of interest to give the expressions for some of the associated Legendre functions corresponding to the expressions for the functions of zero order which were given in equations (31). The first ten of these polynomials (including the functions of zero order) are as follows:

$$P_0^0(x) = 1,$$

$$P_1^0(x) = x; \quad P_1^1(x) = (1-x^2)^{\frac{1}{2}} \cdot 1;$$

$$P_2^0(x) = \tfrac{1}{2}(3x^2 - 1); \; P_2^1(x) = (1-x^2)^{\frac{1}{2}} \cdot 3x; P_2^2(x) = (1-x^2) \cdot 3;$$

$$P_3^0(x) = \tfrac{1}{2}(5x^3 - 3x); \; P_3^1(x) = (1-x^2)^{\frac{1}{2}} \cdot \tfrac{3}{2}(5x^2 - 1);$$

$$P_3^2(x) = (1-x^2) \cdot 15x; \; P_3^3(x) = (1-x^2)^{\frac{3}{2}} \cdot 15, \tag{47}$$

where $x = \cos\theta$.

Since $d^k P_k(x)/dx^k = C$, a constant, it follows that the functions $P_k^k(x)$, for which the order and degree are identical, possess no nodes, but pass through a maximum at $x = 0$ (i.e., $\theta = \pi/2$) as is evident from the relation

$$P_k^k(\cos\theta) = C \cdot \sin^k \theta. \tag{48}$$

[8] It is becoming customary in treatises on quantum mechanics to designate the *square* of the expressions on the right-hand side of (35) and (46) or the *integrals* in (34) and (45) by N. For instance, H. Bethe, in "Handbuch der Physik," Vol. XXIV, Part 1, uses this notation, which may give rise to some confusion for the reader who consults that discussion.

Figure 28[9] shows plots of the four *normalized* functions $P_3^0(\cos\theta)$, $P_3^1(\cos\theta)$, $P_3^2(\cos\theta)$, and $P_3^3(\cos\theta)$ as functions of θ. The normalization factors, as deduced from equation (46), are $N = \sqrt{2/7}$, $\sqrt{24/7}$, $\sqrt{240/7}$, and $\sqrt{1440/7}$, respectively. The actual relations for the

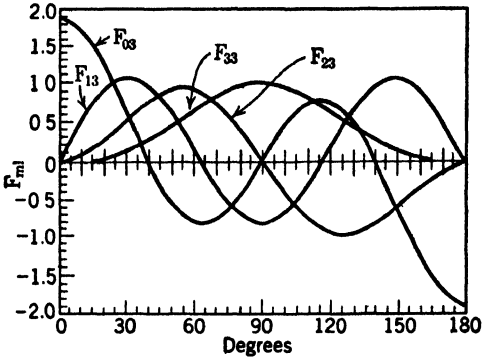


FIG. 28. Plots of the normalized Legendre functions $P_3^m(\cos\theta)$.

four *normalized* functions, and the designations for the corresponding curves, are

$$F_{03} = P_3^0(\cos\theta) = 0.936\ (5\cos^3\theta - 3\cos\theta),$$
$$F_{13} = P_3^1(\cos\theta) = 0.810\sin\theta\ (5\cos^2\theta - 1),$$
$$F_{23} = P_3^2(\cos\theta) = 2.563\sin^2\theta\cdot\cos\theta,$$
$$F_{33} = P_3^3(\cos\theta) = 1.046\sin^3\theta.$$

The nodal points were obtained by solving the equations $P_3^m(\cos\theta) = 0$; the points of maximum values were obtained by solving the equation

$$\frac{dP_3^m(\cos\theta)}{d\theta} = 0,$$

where $m = 0, 1, 2, 3$.

6.7 Geometrical Interpretation of Surface Spherical Harmonics. Let us consider now the geometrical interpretation of the Legendre functions and tesseral harmonics which have been discussed in the previous sections.

The function

$$(A\cos m\eta + B\sin m\eta)P_k^m(\cos\theta) \tag{49}$$

represents a surface spherical harmonic of kth degree and mth order. If $m = 0$, the function has the form $P_k(\cos\theta)$, which is Legendre's

[9] Condon and Morse, "Quantum Mechanics," p. 55.

coefficient of the first kind, of degree k. This function is a polynomial of degree k and therefore has k distinct zero points between $\cos\theta = -1$ and $\cos\theta = +1$. As shown in the curves in Fig. 26 and Fig. 27, these "nodes" are arranged symmetrically about $\cos\theta = 0$, i.e., $\theta = \pi/2$. Hence, on a sphere with the origin as center, the function $P_k(\cos\theta)$ becomes 0 on k different circles, which, as shown in Fig. 29, correspond to different degrees of "latitude," that is, they possess poles at $\theta = 0$ and $\theta = \pi$. These circles are symmetrical with respect to the "equatorial" circle, and if k is odd, the latter is one of the set of circles for which $P_k(\cos\theta) = 0$. Furthermore, as shown by the plots of the functions, since the value of any function $P_k(\cos\theta)$ exhibits $k - 1$ loops, there are $2(k - 1)$ circles parallel to the nodal circles at which

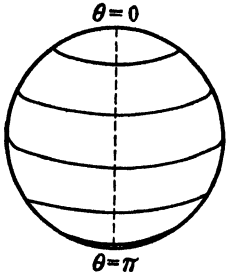


FIG. 29. Geometrical illustration of zonal harmonics.

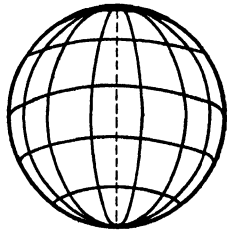


FIG. 30. Geometrical illustration of tesseral harmonics.

the function has the same absolute value. It is for this reason that the Legendre coefficients of zero order are known as *zonal harmonics*. The point $\theta = 0$ is designated the *pole*, and the diameter through the pole, the *axis* of the zonal harmonic.

For m greater than 0 and less than k, the functions are represented by the expression

$$(A\cos m\eta + B\sin m\eta)\sin^m\theta \cdot \frac{d^m P_k(\cos\theta)}{d^m(\cos\theta)}.$$

This may be written in the form

$$\sqrt{A^2 + B^2}\sin(m\eta + \delta)\cdot\sin^m\theta \cdot \frac{d^m P_k(\cos\theta)}{d^m(\cos\theta)}, \tag{50}$$

where $\tan\delta = A/B$. It vanishes for $m\eta = -\delta$ and $m\eta = \pi - \delta$. On a sphere, as shown in Fig. 30, this corresponds to m great circles through the pole $\theta = 0$ (circles of "longitude"), distributed symmetrically, so that the angle between the planes of any two consecutive circles is equal to π/m.

The factor $\sin^m \theta$ is equal to zero only at $\theta = 0$ and $\theta = \pi$. The differential coefficient $d^m P_k(x)/dx^m$ is represented by a function which is the mth derivative of a polynomial of degree k. Thus the highest power of x has the value $k - m$, and the function has $k - m$ zeros on circles with $\theta = 0$ as pole, which are arranged like the corresponding circles in the case of the zonal harmonics. Since the two sets of circles intersect orthogonally, these harmonics are designated *tesseral harmonics.* (See Fig. 30.)

The sum of the $2k + 1$ tesseral harmonics, for which the general expression is given by (49) or (50), is known as a *surface spherical harmonic of degree* k.

For $m = k$, the differential coefficient becomes a constant factor, and the spherical harmonic is of the form

$$\sqrt{A^2 + B^2} \cdot \sin(k\eta + \delta) \sin^k \theta. \tag{51}$$

As pointed out already, this vanishes on k great circles passing through the points $\theta = 0$ and $\theta = \pi$, the angle between the planes of any two consecutive circles being π/k. Since the sphere is thus divided up into $2k$ sectors, as shown in Fig. 31, these functions are known as *sectorial harmonics.*

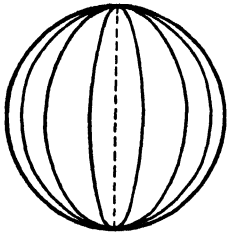


FIG. 31. Geometrical illustration of sectorial harmonics.

For any given value of η, the expression in (51) is evidently of the same form as that in (48), and it is seen that, as the value of k is increased, the function tends to assume appreciable values in an increasingly narrower region symmetrical about the equatorial plane. The interpretation of this result from a physical point of view is considered in the following section.

6.8 The Physical Significance of the Characteristic Functions. We may now consider the significance of the somewhat tedious calculations and seemingly complicated results that have been derived in the previous sections.

The problem to be solved is the following. Given a diatomic molecule, what will be the possible energy states and modes of rotation for such a molecule? The problem first originated because of the observations on the temperature variation of the specific heat of diatomic gases. In order to account for the increase in specific heat with temperature, it was found necessary to assume that, in addition to kinetic energy of translational motion, diatomic molecules also possess an energy due to rotation about an axis of symmetry passing through the center of gravity.

On the basis of the classical quantum theory, the method of determining the discrete energy states of such a system was as follows:

Referring to equation (5), the total energy is given by

$$E = H(p, q) = \frac{1}{2\mu}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\eta^2}{r^2 \sin^2 \theta}\right). \tag{52}$$

In the case of the rigid rotator, $p_r = 0$, and $\mu r_0^2 = I$, the moment of inertia. Hence, the equation for the rotator with free axis is

$$E = \frac{1}{2I}\left(p_\theta^2 + \frac{p_\eta^2}{\sin^2 \theta}\right). \tag{53}$$

Since the evaluation of the integral $\oint p_\theta d\theta$ is rather involved,[10] we shall consider only the case of rotation in a plane. For this case, $p_\theta = 0$ and $\sin^2 \theta = 1$. Hence,

$$2IE = p_\eta^2. \tag{i}$$

From the canonical relation

$$\dot{p}_\eta = -\frac{\partial H}{\partial \eta} = 0,$$

it follows that

$$p_\eta = \alpha, \text{ a constant} \tag{ii}$$

$$= \sqrt{2IE} \text{ from (i).}$$

Therefore,

$$\int_0^{2\pi} p_\eta d\eta = \sqrt{2IE} \cdot 2\pi.$$

Since, in accordance with the Wilson-Sommerfeld quantum condition,

$$2\pi\sqrt{2IE} = mh,$$

where $m = 0, 1, 2, 3$, etc., it follows that

$$E_m = \frac{m^2 h^2}{8\pi^2 I}. \tag{iii}$$

In the case of the rotator with free axis, the calculation leads to the same result, where m is the sum of the two quantum numbers, one corresponding to p_η and the other to p_θ.

[10] Full details will be found in M. Born's "Atommechanik" and Sommerfeld's "Atombau und Spektrallinien."

But observations on band spectra, in which the lines constituting individual bands are due to transitions between states differing in amounts of rotational energy, showed that this result was not quite satisfactory. On the basis of the S. equation, as shown in equation (14), the same result is deduced, if it is assumed that the molecule is capable of rotation in only one plane. The result deduced in equation (25) is, however, in very good agreement with observations on band spectra. Hence, we conclude that diatomic molecules possess two modes of motion about their center of gravity, one in which there is a rotation in the plane containing the axis of the molecule, about an axis of symmetry at right angles to this plane, and another which corresponds to a precession of the axis of the molecule about the axis of symmetry.

Now it is the essence of the S. equation that it starts with this *physical model* and then, instead of discussing the consequences to be deduced from this model by classical mechanics, considers a partial *differential equation* which is derived from the physical model by a definite mathematical procedure, and which represents mathematically the propagation of a wave motion. We abandon, as it were, the concrete, tangible model of a dumbbell-shaped mass rotating about an axis and consider instead the nature and properties of certain wave patterns obtained by solving the partial differential equation.

In the previous sections it was found that corresponding to each energy state of quantum number k, as defined by equation (25), there are $2k + 1$ characteristic functions which represent $2k + 1$ different possible modes of vibration. The next question to be considered is this. What is the physical interpretation of these functions which we recognized as Legendre polynomials of the first kind?

As in the case of the characteristic functions ϕ ("amplitude" functions), deduced in solving the problem of the linear harmonic oscillator, we assign physical interpretations to $\phi\bar{\phi}$ or ϕ^2. In the case of the rigid rotator in three dimensions, the function ϕ is defined by the relation

$$\phi = \frac{1}{\sqrt{2\pi}} \cdot \epsilon^{im\eta} \cdot \frac{P_k^m(x)dx}{N}, \qquad (x = \cos\theta)$$

where N is the normalizing factor for the particular Legendre polynomial. This is determined from the relation

$$\frac{1}{N^2}\int_{-1}^{1} \{P_k^m(x)\}^2 dx = 1. \tag{54}$$

Hence $(1/N)^2 \{P_k^m(\cos\theta)\}^2 \sin\theta d\theta$ is regarded as representing the probability of locating the particle in the region on the surface of a sphere lying between the zonal circles θ and $\theta + d\theta$, while $\left(\frac{1}{2\pi}\right) \epsilon^{im\eta}\epsilon^{-im\eta}\, d\eta =$

$\left(\frac{1}{2\pi}\right) d\eta$, represents the probability of occurrence in the meridian section located between the angles η and $\eta + d\eta$. Evidently this probability is independent of η.

From Figs. 22*a* and 22*b* it follows that the element of area dA on the surface of a sphere of unit radius is given by

$$dA = \sin \theta d\theta d\eta,$$

and the probability of locating the particle in this area at the angles θ and η is

$$PdA = \frac{1}{2\pi N^2} \{P_k(\cos \theta)\}^2 \sin \theta d\theta d\eta. \tag{55}$$

Hence,

$$P = \frac{1}{2\pi} \left\{ \frac{(2k+1)\ (k-m)!}{2(k+m)!} \right\} \{P_k^m\ (\cos \theta)\}^2 \tag{56}$$

corresponds to a *probability per unit area* or " probability density."[11]

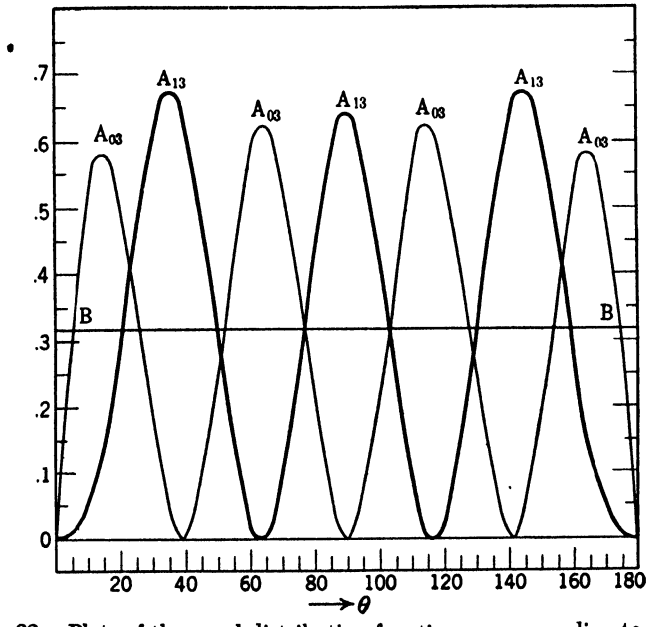


FIG. 32. Plots of the zonal distribution functions corresponding to $k = 3$.

From values of the normalized polynomials as functions of θ, such as illustrated in the plots in Fig. 28, it is possible to calculate both P and $P \sin \theta$, and results obtained in this manner are illustrated by the plots

[11] The function P corresponds to the product $\Theta_{km}(\theta) \cdot \Phi(\eta)$ given by Pauling and Wilson, "Introduction to Quantum Mechanics," pp. 132–133.

shown in Figs. 32, 33, 34, and 35. The first two figures give values of $P \sin \theta$. The designations on the curves and the corresponding *normalized* functions are as follows:

$$A_{03} \ldots\ldots \{P_3^0(\cos\theta)\}^2 \sin\theta.$$
$$A_{13} \ldots\ldots \{P_3^1(\cos\theta)\}^2 \sin\theta.$$
$$A_{23} \ldots\ldots \{P_3^2(\cos\theta)\}^2 \sin\theta.$$
$$A_{33} \ldots\ldots \{P_3^3(\cos\theta)\}^2 \sin\theta.$$

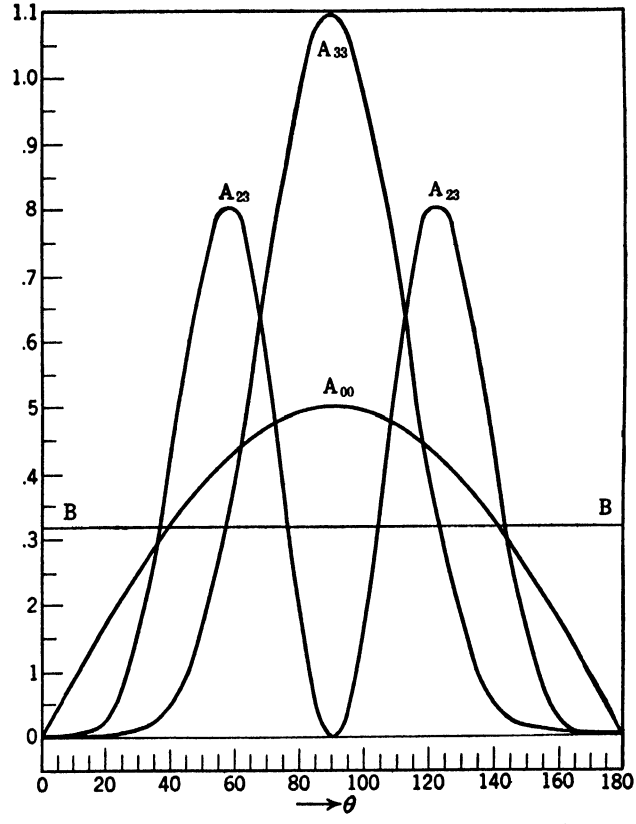


FIG. 33. Plots of the zonal distribution functions corresponding to $k = 3$.

For comparison there has also been plotted, as curve A_{00} in Fig. 33, the function

$$\frac{1}{N^2}\{P_0(\cos\theta)\}^2 \sin\theta = \frac{\sin\theta}{2}.$$

Since

$$\int_0^\pi \sin\theta d\theta = 2$$

and the *total area under each of the curves shown in Fig.* 32 *and Fig.* 33 *is equal to* 1, the average ordinate is given by $1/\pi = 0.318$. This has been indicated by the straight line BB in the two figures.

If now we compare two zones of equal widths, at the angles θ_1 and θ_2, it is evident that the areas of the two zones will be $2\pi \sin \theta_1 d\theta$ and $2\pi \sin \theta_2 d\theta$, where $d\theta$ is the width of each zone, and the radius of the sphere is taken as unity. Hence the relative values of the probability density, as given by P in equations (55) and (56) and plotted in Figs. 34 and 35, are quite different from the values $P \sin \theta$ shown in Figs. 32 and 33. In the former, the distance from the center to any point on the curve gives the relative value of P at the corresponding value of θ. In Fig. 34 the function $\{P_3^0(\cos \theta)\}^2/N^2$ has been plotted and should be compared with the curve A_{03} in Fig. 32, while the plot in Fig. 35 which corresponds to $\{P_3^3(\cos \theta)\}^2/N^2$ is to be compared with the curve A_{33} in Fig. 33.

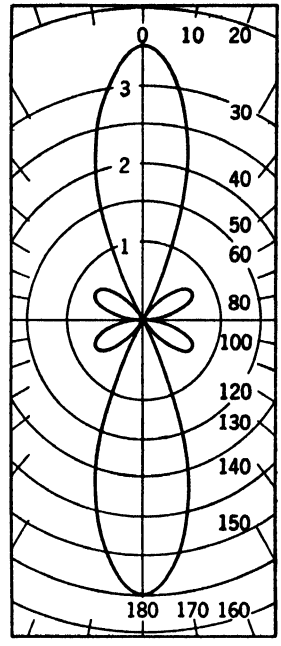


FIG. 34. Probability density function corresponding to associated Legendre function $P_3^0(\cos \theta)$.

In terms of the model of a diatomic molecule these plots indicate that the axis of the molecule will tend to be oriented with respect to the axis of symmetry in those directions for which P is a maximum. This interpretation is most readily evident from the plot A_{33} in Fig. 33 and the corresponding plot in Fig. 35.

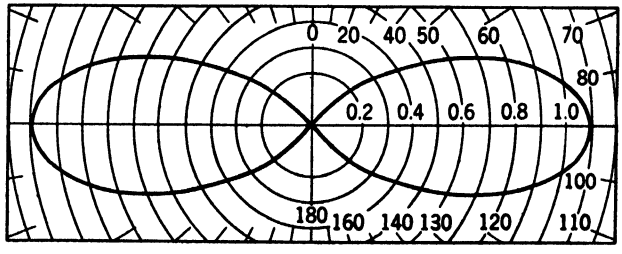


FIG. 35. Probability density function corresponding to associated Legendre function $P_3^3(\cos \theta)$.

In this case there is a relatively narrow region about the value $\theta = \pi/2$ for which P is a maximum. As the value of k is increased (keeping $m = k$), the width of this region decreases rapidly. That is, in the

rotational states of higher energy content the molecule will tend more and more to rotate about an axis of symmetry at right angles to the axis of the molecule.

6.9 Angular Momentum for Motion of Rotator. In Chapter II, it was shown that in case of motion along a coördinate x, the corresponding momentum is obtained by solving the equation

$$p\phi = \frac{h}{2\pi i} \cdot \frac{d\phi}{dx}. \tag{57}$$

If the result of performing the operation on the right-hand side of this equation is of the form $\alpha\phi$, where α is a constant, the conclusion is drawn that this will be the value of the momentum observed in any experiment arranged for this purpose when the particle is in the state designated by the eigenfunction ϕ.

In the case of angular momenta equation (57) is not applicable, as appears from the following consideration.

As shown in section 8, the Hamiltonian form of expression for the energy E is given by

$$E = \frac{1}{2I}\left(p_\theta^2 + \frac{1}{\sin^2\theta} \cdot p_\eta^2\right). \tag{53}$$

If the same rule were followed as that used for converting an expression for E in terms of rectangular coördinates and their corresponding momenta into a S. equation (see Chapter II), the resulting differential equation would be of the form

$$E\phi = -\frac{h^2}{8\pi^2 I}\left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin^2\theta} \cdot \frac{\partial^2}{\partial\eta^2}\right)\phi,$$

that is,

$$\frac{\partial^2\phi}{\partial\theta^2} + \frac{1}{\sin^2\theta} \cdot \frac{\partial^2\phi}{\partial\eta^2} + \frac{8\pi^2 IE}{h^2} \cdot \phi = 0,$$

which obviously is not identical with equation (11), and is not the correct form of S. equation to represent the particular problem.

We will now try to deduce a rule for calculating the angular momenta with respect to θ and η for a rotating body, on the basis of wave mechanics. As a first step we consider the relation which exists according to ordinary mechanics between the angular momentum with respect to the z-axis (the axis from which θ is measured) and the linear momenta p_x and p_y, with respect to the x- and y-axes, respectively. In order to simplify the calculations we shall assume that the motion occurs in the XOY plane only (see Fig. 22a), so that $d\theta/dt = 0$ and $\theta = \pi/2$.

We then have the following relations:

$$x = r\cos\eta; \quad dx = -r\sin\eta d\eta + \cos\eta dr; \tag{i}$$
$$y = r\sin\eta; \quad dy = r\cos\eta d\eta + \sin\eta dr. \tag{ii}$$

Hence,

$$xdy - ydx = (xr\cos\eta + yr\sin\eta)d\eta + (x\sin\eta - y\cos\eta)dr = r^2 d\eta, \tag{iii}$$

since $x^2 + y^2 = r^2$, and the coefficient of dr is equal to zero.

But

$$\frac{\mu r^2 d\eta}{dt} = \mu r^2\dot{\eta} = M_z,$$

where M_z denotes the angular momentum with respect to the z-axis. Therefore we have the relation

$$M_z = \mu(x\dot{y} - y\dot{x}) = xp_y - yp_x, \tag{58}$$

which is of extreme importance in classical mechanics, since it enables us to express angular momentum in terms of the rectangular coördinates and their associated momenta.

Now in deriving the S. equation from the expression for the energy in terms of rectangular coördinates and associated momenta, we set

$$p_x = \frac{h}{2\pi i}\cdot\frac{\partial}{\partial x} \quad \text{and} \quad p_y = \frac{h}{2\pi i}\cdot\frac{\partial}{\partial y}.$$

We therefore conclude that in quantum mechanics we are justified in assuming that M_z may be used *as an operator*, which is defined by the relation

$$M_z = \frac{h}{2\pi i}\left(x\cdot\frac{\partial}{\partial y} - y\cdot\frac{\partial}{\partial x}\right). \tag{59}$$

Now let us assume that we have any function F which is a function of the polar coördinates η and r, or of the coördinates x and y. We have the following relations between the differential coefficients:

$$\frac{\partial F}{\partial \eta} = \frac{\partial F}{\partial x}\cdot\frac{dx}{d\eta} + \frac{\partial F}{\partial y}\cdot\frac{dy}{d\eta},$$

and

$$\frac{\partial F}{\partial r} = \frac{\partial F}{\partial x}\cdot\frac{dx}{dr} + \frac{\partial F}{\partial y}\cdot\frac{dy}{dr}.$$

Substituting from (i) and (ii) it follows that

$$\frac{\partial F}{\partial \eta} = x\cdot\frac{\partial F}{\partial y} - y\cdot\frac{\partial F}{\partial x} = \left(x\cdot\frac{\partial}{\partial y} - y\cdot\frac{\partial}{\partial x}\right)F. \tag{iv}$$

Hence we conclude, by comparing (iv) with (59), that *as an operator*

$$M_z = \frac{h}{2\pi i} \cdot \frac{\partial}{\partial \eta} \cdot \tag{60}$$

Thus, in the case of the rotator with fixed axis, the normalized function, as given by equation (19), is

$$Z_{\pm m} = \frac{\epsilon^{\pm im\eta}}{\sqrt{2\pi}} \cdot$$

Therefore, in order to determine whether the angular momentum has a definite value, we consider the equation

$$M_z Z = \frac{h}{2\pi i} \cdot \frac{\partial Z}{\partial \eta} = \frac{h}{2\pi i} \left(\frac{\pm im}{\sqrt{2\pi}} \right) \epsilon^{\pm im\eta} = \pm \frac{hm}{2\pi} \cdot Z. \tag{61}$$

That is, M_z operating on the function Z, yields as a result a constant multiplied by Z. Hence, we conclude that an experiment arranged to determine the magnitude and sign of the angular momentum would lead to a value $\pm mh/(2\pi)$, depending on the relation between the direction of rotation and that of the perturbing field. Here the result deduced by the operator method is identical with that deduced by ordinary mechanics, when the quantizing principle is introduced. As stated previously, the observations on band spectra show that it is not correct to treat the problem of a rotating molecule as one in only two dimensions. It is therefore necessary to determine what the form of the operator must be for the case of a rotator in three dimensions.

This problem is somewhat more complicated because, as is evident from equation (53), the angular momentum terms for θ and η do not enter into the expression for E in the same manner. It turns out[12] that under these circumstances it is most convenient to calculate the square of the resultant angular momentum vector $\mathbf{M}^2$, which is defined thus:

$$\mathbf{M}^2 = \mathbf{M}_x^2 + \mathbf{M}_y^2 + \mathbf{M}_z^2,$$

and as an operator, this is defined by the relation

$$M^2 = -\frac{h^2}{4\pi^2} \left[\left(y \cdot \frac{\partial}{\partial z} - z \cdot \frac{\partial}{\partial y} \right)^2 + \left(z \cdot \frac{\partial}{\partial x} - x \cdot \frac{\partial}{\partial z} \right)^2 + \left(x \cdot \frac{\partial}{\partial y} - y \cdot \frac{\partial}{\partial x} \right)^2 \right].$$

[12] This is discussed more fully by E. C. Kemble, *Phys. Rev., Suppl.*, **1**, 157 (1929); also see J. Frenkel, "Wellenmechanik," pp. 248–253 (1929 edition).

Introducing spherical coördinates, and proceeding as in the case of the single variable η, it may be shown that the differential equation for the determination of M^2 is of the form

$$\frac{h^2}{4\pi^2}\left[\frac{1}{\sin\theta}\cdot\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial\phi}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\cdot\frac{\partial^2\phi}{\partial\eta^2}\right]+M^2\phi=0. \tag{62}$$

That is, M^2 *as an operator* has the form

$$M^2=-\frac{h^2}{4\pi^2}\left[\frac{1}{\sin\theta}\cdot\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\cdot\frac{\partial^2}{\partial\eta^2}\right]=-\frac{h^2}{4\pi^2}\cdot\Omega,$$

where Ω is an operator of the same type as the Laplacian operator ∇^2. In fact, equation (4) may be written in the form

$$\nabla^2\phi=\frac{1}{r^2}\Omega\phi+\frac{1}{r^2}\left(2r\cdot\frac{\partial\phi}{\partial r}+r^2\cdot\frac{\partial^2\phi}{\partial r^2}\right)=\frac{1}{r^2}\Omega\phi+\frac{1}{r}\cdot\frac{\partial^2(r\phi)}{\partial r^2}.$$

Equation (62) may therefore be written in the form

$$\left(M^2+\frac{h^2}{4\pi^2}\Omega\right)\phi=0. \tag{63}$$

Similarly, equation (11) may be written as

$$\left(2EI+\frac{h^2}{4\pi^2}\Omega\right)\phi=0,$$

the solution of which, as shown previously, is

$$2EI=\frac{k(k+1)h^2}{4\pi^2}.$$

Hence, the solution of (63) must also be the same. That is,

$$M^2=\frac{k(k+1)h^2}{4\pi^2}, \tag{64}$$

while, as deduced in equation (61),

$$M_z=\pm\frac{hm}{2\pi}. \tag{65}$$

Equation (64) leads to the conclusion that the *total angular momentum* of the rotating system may be designated by a vector whose magnitude is

$$\mathbf{M}=\frac{h}{2\pi}\sqrt{k(k+1)},$$

and that the *component* of this vector along the z-axis is given by $mh/2\pi$. In terms of a *unit vector* of magnitude $h/(2\pi)$, the total angular momentum and the component of this vector along the z-axis are therefore $\sqrt{k(k+1)}$ and m, respectively.

These results are specially significant in connection with the problem of the hydrogen atom, which is discussed in the following chapter.

SUPPLEMENTARY NOTE 1

EXPANSION OF AN ARBITRARY FUNCTION IN TERMS OF AN ORTHOGONAL SYSTEM OF FUNCTIONS

In subsequent discussions we shall have occasion to make use of the very important property of orthonormalized functions which is expressed in the form

$$\int \phi_n \bar{\phi}_m dv = 0 \quad (m \neq n)$$
$$= N^2 \ (m = n),$$

where ϕ_n and ϕ_m are any two eigenfunctions of the system, N is the normalizing factor, dv is the element of volume, area, or length, and the integration is extended over the whole region in which the functions are physically significant. (This region is usually designated the "configuration space.")

As has been mentioned previously, the simplest type of such expressions is the trigonometric functions

$$\frac{1}{\sqrt{\pi}} \sin m\theta \quad \text{and} \quad \frac{1}{\sqrt{\pi}} \cos n\theta,$$

for which the limits are 0 and 2π, so that

$$\int_0^{2\pi} \cos m\theta \cos n\theta d\theta = \int_0^{2\pi} \sin m\theta \sin n\theta d\theta = \begin{matrix} 0 & (m \neq n) \\ \pi & (m = n) \end{matrix} \Big\}. \qquad \text{(i)}$$

By means of this relation it becomes possible to develop any given function of θ in terms of the sines or of the cosines of multiples of θ. Series involving only these trigonometric functions, that is, series of the form

$$a_0 + a_1 \cos x + a_2 \cos 2x + \ldots + a_n \cos nx + \ldots$$
$$b_1 \sin x + b_2 \sin 2x + \ldots + b_m \sin mx + \ldots$$

are known as *Fourier's series.* The possibility of expressing any arbitrary function of θ in terms of such a series may be illustrated by the following examples.

(1) It is required to develop the function $f(\theta) = \theta$ as a sine series for the region $\theta = 0$ to $\theta = \pi$. We have

$$\theta = a_1 \sin \theta + a_2 \sin 2\theta + \ldots + a_n \sin n\theta + \ldots$$

Let us multiply each side of this equation by $\sin n\theta d\theta$ and integrate between 0 and π. Then

$$\int_0^\pi \theta \sin n\theta d\theta = a_1 \int_0^\pi \sin\theta \sin n\theta d\theta + \ldots + a_n \int_0^\pi \sin^2 n\theta d\theta + \ldots$$

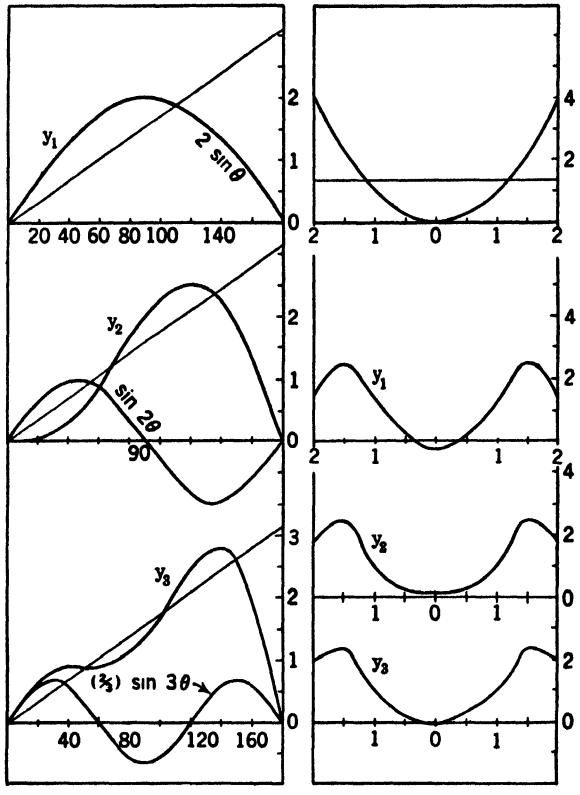


FIG. 36. Illustrating Fourier's series analysis.

Because of equation (i) all the terms on the right-hand side, except the one involving $\sin^2 n\theta$, vanish. Hence we obtain a relation for determining a_n, which is of the form

$$a_n = \frac{2}{\pi} \int_0^\pi \theta \sin n\theta d\theta. \qquad \text{(ii)}$$

Now

$$d(\theta \cdot \cos n\theta) = -n\theta \cdot \sin n\theta d\theta + \cos n\theta \cdot d\theta$$

$$= -n\theta \cdot \sin n\theta d\theta + \frac{1}{n} \cdot d(\sin n\theta).$$

Hence,

$$\int_0^\pi \theta \sin n\theta d\theta = -\frac{1}{n} \cdot \theta \cos n\theta \Big]_0^\pi + \frac{1}{n^2} \sin n\theta \Big]_0^\pi = \frac{(-1)^{n+1}\pi}{n},$$

since $\sin n\theta = 0$ for both $\theta = 0$ and $\theta = \pi$, while $\cos n\pi = (-1)^n$.

In a similar manner all the other coefficients, a_1, a_2, etc., may be determined and the required development has the form

$$\theta = 2\left(\frac{\sin\theta}{1} - \frac{\sin 2\theta}{2} + \frac{\sin 3\theta}{3} - \ldots\right). \tag{iii}$$

Figure 36 shows in the left-hand series of plots, the straight line $y = \theta$ from $\theta = 0$ to $\theta = \pi$ (that is from 0 to 180 degrees), and the successive approximations to this line which are obtained by taking

$$y_1 = 2\sin\theta;$$
$$y_2 = 2\sin\theta - \sin 2\theta;$$
$$y_3 = 2\sin\theta - \sin 2\theta + (\tfrac{2}{3})\sin 3\theta.$$

It will be observed that while the series does not converge very rapidly, the curves gradually approximate $y = \theta$ more and more closely, with increase in number of terms.

(2) It is desired to express x^2 as a cosine function for the range $x = -c$ to $x = c$.

We introduce a new variable, $z = \pi x/c$, so that $z = -\pi$ for $x = -c$, and $z = \pi$ for $x = c$. Then,

$$x^2 = \left(\frac{c}{\pi} \cdot z\right)^2 = a_0 + a_1 \cos z + a_2 \cos 2z + \ldots \tag{iv}$$

Multiplying each side by $\cos nzdz$ and integrating between the limits, all coefficients on the right-hand side, except $a_n \int_{-\pi}^{\pi} \cos^2 nzdz$, vanish, and we obtain the relation for determining a_n, of the form

$$\frac{c^2}{\pi^2}\int_{-\pi}^{\pi} z^2 \cos nzdz = a_n \int_{-\pi}^{\pi} \cos^2 nzdz = a_n \cdot \pi,$$

that is,

$$a_n = \frac{c^2}{\pi^3}\int_{-\pi}^{\pi} z^2 \cos nz \cdot dz. \tag{v}$$

Now
$$d(z^2 \sin nz) = nz^2 \cos nzdz + 2z \sin nzdz;$$
$$d(z \cos nz) = -nz \sin nzdz + \cos nzdz.$$

Hence,

$$\int_{-\pi}^{\pi} z^2 \cos nz dz = \frac{z^2}{n} \cdot \sin nz \Big]_{-\pi}^{\pi} - \frac{2}{n}\int_{-\pi}^{\pi} z \sin nz dz$$

$$= -\frac{2}{n}\int_{-\pi}^{\pi} z \sin nz dz \text{ (since } \sin n\pi = 0)$$

$$= \frac{2}{n^2}\left[z \cos nz \right]_{-\pi}^{\pi} - \frac{1}{n^2}\int_{-\pi}^{\pi} \cos nz dz$$

$$= \frac{4\pi(-1)^n}{n^2} - \frac{1}{n^2} \cdot \frac{\sin nz}{n} \Big]_{-\pi}^{\pi} = \frac{4\pi(-1)^n}{n^2}. \qquad \text{(vi)}$$

Therefore,

$$a_n = \frac{c^2 \cdot 4\pi(-1)^n}{\pi^3 n^2} = \frac{4c^2(-1)^n}{\pi^2 n^2}. \qquad \text{(vii)}$$

The coefficient a_0 is determined from the relation

$$\frac{c^2}{\pi^2}\int_{-\pi}^{\pi} z^2 dz = \int_{-\pi}^{\pi} a_0 dz, \text{ since } \cos 0 = 1.$$

That is,

$$a_0 \cdot 2\pi = \frac{2c^2\pi^3}{3\pi^2}; \quad a_0 = \frac{c^2}{3} = \frac{1}{c}\int_0^c x^2 dx.$$

Consequently, the development for x^2 is of the form

$$x^2 = \frac{c^2}{3} - \frac{4c^2}{\pi^2}\left(\cos\frac{\pi x}{c} - \frac{1}{2^2}\cos\frac{2\pi x}{c} + \frac{1}{3^2}\cos\frac{3\pi x}{c} - \ldots \right). \qquad \text{(viii)}$$

The curves on the right-hand side of Fig. 36 show the parabola $y = x^2$ at the top, and the straight line $y = c^2/3 = 2^2/3$, which is evidently the average value of y over the range $0 \leqslant |x| \leqslant |2|$. The other curves correspond to the expressions:

$$y_1 = \frac{c^2}{3} - \frac{4c^2}{\pi^2} \cdot \cos\frac{\pi x}{c};$$

$$y_2 = y_1 + \frac{c^2}{\pi^2} \cdot \cos\frac{2\pi x}{c};$$

$$y_3 = y_2 - \frac{4}{9} \cdot \frac{c^2}{\pi^2} \cdot \cos\frac{3\pi x}{c}.$$

As shown by the plot of y_3, this expression corresponds fairly closely to $y = x^2$ for the range $x = \pm 0.75 \cdot c$, and, by using more terms, the range over which the series represents the parabolic function may be made to approach the limits $x = \pm c$ very satisfactorily.

More generally, any function $f(x)$ can be expressed within a definite range of values of x in the form

$$f(x) = a_0 + \sum_1^\infty a_n \cos nx + \sum_1^\infty b_n \sin nx,$$

where

$$a_n = \frac{1}{\pi}\int_{-c}^{c} f(x) \cdot \cos\frac{n\pi x}{c} \cdot \frac{\pi}{c}\,dx;$$

$$b_n = \frac{1}{\pi}\int_{-c}^{c} f(x) \cdot \sin\frac{n\pi x}{c} \cdot \frac{\pi}{c}\,dx.$$

Such a representation in terms of trigonometric functions is known as a " Fourier's series " expansion for $f(x)$.

The possibility of obtaining such developments of an arbitrary function depends, evidently, upon the existence of the orthogonality relation expressed by equation (i). The same type of reasoning may be applied to develop an arbitrary function of x between the limits 0 and ∞ in terms of a series of Hermitian or Laguerre polynomials.

Thus if $f(x)$ is a function which tends to vanish for $x = \pm\infty$, we can obtain the coefficients a_n in the series

$$f(x) = \epsilon^{-\frac{x^2}{2}}\{a_0H_0(x) + a_1H_1(x) \ldots a_nH_n(x) + \ldots\}$$

from the relation

$$a_n = \frac{\int_0^\infty f(x) \cdot \epsilon^{-\frac{x^2}{2}} \cdot H_n(x)dx}{\int_0^\infty \epsilon^{-x^2} \cdot H_n^2(x)dx}$$

$$= 2^{n-1}n!\sqrt{\pi}\int_0^\infty f(x)\epsilon^{-\frac{x^2}{2}} \cdot H_n(x)dx. \qquad \text{(ix)}$$

Similarly, $f(\theta)$, an arbitrary[13] function of θ, may be represented in the range $\pi > \theta > 0$ by the series of Legendre coefficients of zero order, in the form

$$f(\theta) = A_0P_0(\cos\theta) + A_1P_1(\cos\theta) + \ldots + A_nP_n(\cos\theta) + \ldots$$

[13] "Arbitrary" in the sense that it is possible to plot the function graphically.

where

$$A_n = \frac{\int_0^\pi f(\theta) \cdot P_n(\cos\theta) \cdot \sin\theta d\theta}{\int_0^\pi \{P_n(\cos\theta)\}^2 \sin\theta d\theta}$$

$$= \frac{2n+1}{2} \int_0^\pi f(\theta) \cdot P_n(\cos\theta) \cdot \sin\theta d\theta. \qquad \text{(x)}$$

Even if the integral in equation (ix) or (x) cannot be calculated by direct integration, it can always be evaluated by plotting the integrand (that is, the expression to be integrated) as a function of x or θ and determining the area under the plot graphically. However, it is usually possible for the experienced mathematician to develop a convergent series for the integral, by means of which its actual value may be determined.

More generally, if an arbitrary function ψ can be developed in terms of an *orthonormalized* set of functions $\phi_0, \phi_1, \ldots \phi_n$, such that

$$\psi = \sum_0^n a_n \phi_n, \qquad \text{(xi)}$$

then

$$a_n = \int \psi \phi_n dv, \qquad \text{(xii)}$$

where the integration is carried out over the configuration space.

In case ψ is a complex function, then ϕ_n must be replaced by $\bar{\phi}_n$, so that a_n is real.

COLLATERAL READING

Mathematical Treatises on Fourier's Series and Spherical Harmonics

1. Byerly, W. E., "Fourier's Series and Spherical Harmonics," Ginn & Co., Boston. A standard treatise for reference.
2. MacRobert, T. M., "Spherical Harmonics," E. P. Dutton & Co., New York. A more elementary treatment than Byerly's.
3. Humphrey, D., "Advanced Mathematics," Chapter II.
4. Ingersoll, L. R., and Zobel, O. J., "An Introduction to the Mathematical Theory of Heat Conduction," Ginn & Co., Boston, 1913. This is an elementary work, containing numerous illustrations of the application of Fourier's series.
5. Houstoun, R. A., "An Introduction to Mathematical Physics," also contains an excellent discussion of Fourier's series.
6. Jeans, J. H., "Electricity and Magnetism," Cambridge Univ. Press. This treatise also contains a very comprehensive treatment of spherical harmonics.
7. Kellogg, O. D., "Foundations of Potential Theory," Julius Springer, Berlin, 1929. Excellent discussion of Legendre functions.

8. COURANT, R., and HILBERT, D., "Methoden der Mathematischen Physik, I," Chapter II, p. 67 et seq.
9. SILBERSTEIN, L., "Mathematical Tables," Chapter II, p. 132 et seq.
10. JAHNCKE, E., and EMDE, F., "Funktionentafeln," B. G. Teubner, Berlin, 1928.

Problem of Rigid Rotator

1. SCHROEDINGER, E., *Ann. Physik*, **79**, 520 (1926).
2. SOMMERFELD, A., "Wave Mechanics"; contains comprehensive treatment of the solutions of the Legendre equation.
3. PAULING, L., and WILSON, E. B., Jr., "Quantum Mechanics." Chapter V treats of the topic also very fully. On pp. 148–149 there are given a number of plots similar to Figs. 31 and 32.
4. KEMBLE, E. C., *Phys. Rev. Suppl.*, **1**, 157 (1929), gives a discussion of the momentum operator and of the expansion problem in terms of orthogonal functions.
5. SOMMERFELD, A., "Atombau und Spektrallinien," F. Vieweg & Sohn Akt.-Ges., Braunschweig (1924). Gives the classical treatment of a number of the problems discussed in this and subsequent chapters.

CHAPTER VII

THE HYDROGEN ATOM

7.1 Bohr Theory of the Hydrogen-Like Atom. In Chapter IV it was shown that, on the basis of the Bohr theory, the discrete energy states of the hydrogen-like atom (of nuclear charge $+Ze$) are determined, in accordance with the Wilson-Sommerfeld quantum conditions, by the relation

$$E_n = -\frac{2\pi^2\mu Z^2e^4}{n^2h^2} \tag{1}$$

$$= -\frac{RchZ^2}{n^2}, \tag{2}$$

where $R = \dfrac{2\pi^2e^4\mu}{h^3c}$ is known as the Rydberg constant.[1]

Furthermore, it was shown that, for any given value of the *total quantum number* n, the number of possible electronic orbits is determined by the series of values for the *azimuthal quantum number* $k = n$, $n - 1, \ldots 1$.

For any state of quantum number n, the semi-major axis is given by the relation

$$a_n = -\frac{Ze^2}{2E_n} = \frac{n^2h^2}{4\pi^2\mu Ze^2}$$

$$= n^2a_1, \tag{3a}$$

where

$$a_1 = \frac{h^2}{4\pi^2\mu e^2Z} \tag{3b}$$

is the radius of the Bohr orbit for the state $n = 1$. For the case $Z = 1$, $n = 1$, that is, the *normal state of the hydrogen atom*, the radius is

[1] It is customary to designate this particular value of the Rydberg constant by R_∞. As pointed out in Chapter IV, it is necessary in calculating E_n to take into account the motion of the nucleus as well as that of the electron about their common center of gravity. In that case μ corresponds to the "reduced mass," which becomes equal to that of the electron for infinite nuclear mass.

designated by

$$a_0 = a_1 Z = \frac{h^2}{4\pi^2 \mu e^2}, \tag{3c}$$

and is known as the radius of the normal Bohr orbit.

Substituting the values for the constants in equations (2) and (3c), the results obtained are

$$a_0 = 0.5282 \times 10^{-8} \text{ cm.},$$

and

$$R_\infty = 109737.42 \text{ cm.}^{-1}$$

7.2 Hydrogen Atom as a Potential Barrier Problem. Before discussing the solution of the appropriate S. equation for this problem, it is of interest to point out a method by which an approximate solution may be derived for the values of the series of discrete energy states.

As N. F. Mott has pointed out,[2] "a hydrogen atom is simply an electron bound by an electrostatic force, which pulls it back if it tries to get away from the nucleus. The wave function therefore will vibrate in normal modes, and we shall only have wave functions describing the behavior of electrons of certain discrete energies."

An approximate calculation of the magnitude of these energy states may be made as follows:

It was shown in Chapter III that, in the case of an electron in a potential "box," the series of discrete energy states is given by

$$E_m = \left(\frac{m}{2a}\right)^2 \cdot \frac{h^2}{8\mu}, \tag{3.29}$$

where m is an integral value and $2a$ is the extent of the region between the barriers. Also it was shown that, for the lowest of these states ($m = 1$), the corresponding characteristic function is

$$\phi = 2\sqrt{\frac{1}{\lambda}} \cdot \cos \frac{2\pi x}{\lambda}, \tag{3.34}$$

where

$$\lambda = \frac{h}{\sqrt{2\mu E_1}} = 4a. \tag{3.32}$$

In the case of a hydrogen atom, the field of force is defined by the potential energy $-Ze^2/r$. Since the total energy cannot exceed this value for any given value of r, it follows, if $2a$ denote the diameter of the

[2] N. F. Mott, "An Outline of Wave Mechanics," p. 63.

electron orbit for the lowest state, that

$$E_1 = -\frac{Ze^2}{a} \tag{i}$$

and

$$\lambda = 4a = \frac{h}{\sqrt{-2\mu E_1}}. \tag{ii}$$

From (i) and (ii) it follows, by eliminating a, that

$$E_1 = -\frac{32\mu Z^2 e^4}{h^2},$$

which differs from the expression in (1) by the ratio $32/(2\pi^2)$. If λ had been assumed equal to πa instead of $4a$, the two results would be identical.

7.3 The Schroedinger Equation for the Hydrogen-Like Atom. In the Schroedinger method of calculating energy states the starting point is the same as in the classical method, that is, we consider a system consisting of nucleus of charge $+Ze$, and of an electron for which the potential energy is given by $V = -Ze^2/r$, and the kinetic energy T, by $\frac{1}{2}\mu v^2$. In terms of spherical polar coördinates, the total energy is given by

$$E = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\dot{\eta}^2) - \frac{Ze^2}{r}. \tag{4}$$

As shown in equation (6.8), the S. equation deduced from this relation is of the form

$$\nabla^2\phi + \alpha^2\left(E + \frac{Ze^2}{r}\right)\phi = 0, \tag{5}$$

where $\alpha^2 = 8\pi^2\mu/h^2$, and the Laplacian operator is given by the relation

$$\nabla^2\phi = \frac{1}{r^2 \sin\theta}\left\{\frac{\partial}{\partial r}\left(r^2 \sin\theta \cdot \frac{\partial\phi}{\partial r}\right)\right.$$

$$\left. + \frac{\partial}{\partial\theta}\left(\sin\theta \cdot \frac{\partial\phi}{\partial\theta}\right) + \frac{\partial}{\partial\eta}\left(\frac{1}{\sin\theta} \cdot \frac{\partial\phi}{\partial\eta}\right)\right\}. \tag{6.4}$$

The solution of this partial differential equation must yield $\phi = \phi(r, \theta, \eta)$, and, as in the case of the rigid rotator, we assume that it is possible to express this solution in the form

$$\phi(r, \theta, \eta) = S(r) \cdot Y(\theta, \eta),$$

where $S(r)$ is a function of the radius only, and $Y(\theta, \eta)$ denotes a function of the angle variables.

We thus separate equation (5) into *two* differential equations, one in r, and the other in θ and η, which are as follows:[3]

$$\frac{1}{\sin\theta}\cdot\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial Y}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\cdot\frac{\partial^2 Y}{\partial\eta^2}+CY=0, \tag{6}$$

and

$$\frac{1}{r^2}\cdot\frac{d}{dr}\left(r^2\frac{dS}{dr}\right)-\frac{CS}{r^2}+\alpha^2\left(E+\frac{Ze^2}{r}\right)S=0, \tag{7}$$

where C is an arbitrary constant which plays the same rôle in the present case as the constant m^2, which was introduced in solving equation (*6*.11).

Equation (6) is evidently identical with equation (*6*.11), and the solution is therefore a *tesseral harmonic* of the form[4]

$$Y(\theta,\eta)=P_l^m(\cos\theta)\cdot\epsilon^{im\eta}.$$

$$C=l(l+1),^5$$

where

$$l=0,\ 1,\ 2,\ \text{etc.}$$

and

$$m=\pm l,\ \pm(l-1)\ \ldots\ \pm 1,\ 0.$$

As pointed out in the case of the rigid rotator this signifies that the particular energy state corresponding to any given value E_l is *degenerate*, inasmuch as it may be represented by any one of $2l+1$ independent eigenfunctions. However, it will be observed that it is not possible from the solution of equation (6) to determine the value of E_l. For this purpose it is necessary to solve the radial equation (7), which, by substituting the value for C, takes the form,

$$\frac{d^2S}{dr^2}+\frac{2}{r}\cdot\frac{dS}{dr}+\left\{A+\frac{2B}{r}-\frac{l(l+1)}{r^2}\right\}S=0, \tag{8}$$

where

$$A=\frac{8\pi^2\mu E}{h^2}$$

and

$$B=\frac{4\pi^2\mu Ze^2}{h^2}. \tag{9}$$

[3] Compare section *6*.4.

[4] The tesseral harmonics are usually indicated by the symbol $Y_{l,m}$ where l and m are the two integers used to designate the associated Legendre functions.

[5] The symbol l is introduced instead of k which was used in the previous case, in order to bring the notation into agreement with spectroscopic usage.

7.4 The Radial Schroedinger Equation. Laguerre Functions. In solving the radial equation, we must be guided by the conclusions deduced in ordinary mechanics about the motion of a body in a central field of force. The best illustration of this is furnished by the investigations on the possible orbits of the planets in the gravitational field of the sun. We know that in this case two types of orbits are possible, (1) hyperbolic orbits for which $E > 0$, and (2) elliptical (including circular) orbits, for which $E < 0$. In the quantum mechanics problem we must therefore seek solutions corresponding to these two cases.

Let us consider first the case $E > 0$. For very large values of r, all the terms in (8) involving $1/r$ or $1/r^2$ may be neglected, and the equation becomes

$$\frac{d^2S}{dr^2} + AS = 0. \tag{10}$$

The solution of this equation is evidently

$$S = K_1\epsilon^{ix\sqrt{A}} + K_2\epsilon^{-ix\sqrt{A}}. \tag{11}$$

According to (9),

$$\sqrt{A} = \frac{2\pi\sqrt{2\mu E}}{h} = \frac{2\pi}{\lambda}, \tag{12}$$

where λ is the de Broglie wave length for a particle of kinetic energy E. From the discussion in Chapter II, it is seen that equation (11) represents a combination of two streams, in one of which the particle is receding from the origin, and in the other it is approaching the origin with a constant momentum $\pm\sqrt{2\mu E}$. Obviously E may vary continuously from 0 to *any positive value*. That is, there are no discrete energy states, and this gives an interpretation of the fact that beyond the limits of the line spectrum (which is due to a series of discrete states of negative energy) there is also observed a continuous spectrum, which must correspond to states for which $E > 0$.

The much more interesting case is that for which $E < 0$, and for which, according to the classical Bohr theory, the system exhibits a series of electronic orbits, each of which is distinguished by definite quantum numbers and definite values of the orbital constants.[6]

Since E is negative, $-A$ in equation (9) is positive, and it follows

[6] The following discussion of the problem is based on that of M. V. Laue, "Handbuch der Radiologie," Vol. VI, Part I, pp. 42–48. An elementary presentation is also given by N. F. Mott, "An Outline of Wave Mechanics," pp. 71–77.

from equation (12) that the magnitude

$$a = \frac{1}{2\sqrt{-A}}, \tag{13}$$

which is real, has the dimensions of a length.

We therefore introduce the dimensionless variable,

$$x = \frac{r}{a} = 2r\sqrt{-A}, \tag{14}$$

and consequently equation (8) becomes

$$\frac{d^2S}{dx^2} + \frac{2}{x}\cdot\frac{dS}{dx} + \left\{-\frac{1}{4} + \frac{2Ba}{x} - \frac{l(l+1)}{x^2}\right\}S = 0. \tag{15}$$

The problem is to find solutions of this equation which are finite and continuous for all values of x and which vanish at $x = \infty$. The point $x = 0$ is a singular point, since $1/x$ becomes infinite there. It is therefore necessary to investigate the behavior of the solutions at this point and, in order to do this, we assume that it is possible to express the function S in the form,

$$S = x^s v, \tag{16}$$

where s is a constant and v may be expressed as a polynomial in powers of x, of the form

$$v = \sum_0^\infty v_n x^n.$$

Hence $$S = v_0x^s + v_1x^{s+1} + \ldots + v_nx^{s+n},$$

and $$\frac{dS}{dx} = sv_0x^{s-1} + (s+1)v_1x^s + \ldots,$$

while $$\frac{d^2S}{dx^2} = s(s-1)v_0x^{s-2} + (s+1)sv_1x^{s-1} + \ldots$$

Substituting in equation (15) it is found that the coefficient of v_0x^{s-2} is given by

$$s(s-1) + 2s - l(l+1) = s(s+1) - l(l+1).$$

As we approach $x = 0$, all terms involving higher powers of x than $s - 2$ may be neglected, and if S is to be finite at $x = 0$, it is necessary that $x^{s-2} = 1$; that is, the series must not have any lower powers of x than those given by the relation

$$s(s+1) = l(l+1).$$

It follows that $s = l$, or $s = -(l+1)$, and since x^{-l-1} would become infinite at $x = 0$, we must choose the value $s = l$. Using this value in equation (16), substituting in (15), and dividing by x^l, we obtain the equation

$$\frac{d^2v}{dx^2} + \frac{2(l+1)}{x} \cdot \frac{dv}{dx} + \left(\frac{2Ba}{x} - \frac{1}{4}\right)v = 0. \tag{17}$$

For very large values of x, all terms involving $1/x$ may be neglected, and the equation becomes

$$\frac{d^2v}{dx^2} - \frac{v}{4} = 0,$$

the solution of which is

$$v = C_1\epsilon^{-\frac{x}{2}} + C_2\epsilon^{\frac{x}{2}}.$$

Since v must not increase indefinitely with increase in x (which is always positive), it follows that v should be of the form

$$v = \epsilon^{-\frac{x}{2}} \cdot g(x). \tag{18}$$

Substituting this function and its differential coefficients in (17) and multiplying by $\epsilon^{x/2}$, the result is

$$\frac{d^2g}{dx^2} + \left\{\frac{2(l+1)}{x} - 1\right\}\frac{dg}{dx} + \left\{\frac{2Ba}{x} - \frac{(l+1)}{x}\right\}g = 0. \tag{19}$$

If we examine the behavior of the series

$$g(x) = \sum_{k=0}^{\infty} c_k x^k$$

for large values of x in a manner similar to that used in the case of the Hermitian and Legendre polynomials, it is found that v will decrease exponentially and ultimately vanish for $x \to \infty$ only if the series is made to end with the term in x^j, where j is determined by the relation

$$2Ba = j + l + 1.$$

Substituting for B and a from equations (9) and (13) it follows from the last equation that

$$E = -\frac{2\pi^2\mu Z^2e^4}{h^2} \cdot \frac{1}{(j+l+1)^2}. \tag{20}$$

That is, solutions of the S. equation which are finite and continuous

for all values of x, and which vanish for $x = \infty$, are obtained only for the series of discrete eigenvalues defined by the relation

$$E_n = -\frac{2\pi^2\mu Z^2e^4}{n^2h^2}, \tag{21}$$

where $$n = j + l + 1 = 2Ba$$

and $$j \geqslant 0, l \geqslant 0, \quad \text{while} \quad n \geqslant 1.$$

Equation (21) is identical with equation (1) derived by the Bohr theory, and is in satisfactory agreement with observations on the energy levels for hydrogen-like atoms.

Let us now consider the differential equation (17) which has the form

$$\frac{xd^2v}{dx^2} + 2(l+1)\frac{dv}{dx} + \left(n - \frac{x}{4}\right)v = 0. \tag{22}$$

Let

$$p = 2l + 1$$

and

$$k = n + l.$$

Then

$$\begin{aligned} n &= \tfrac{1}{2}(2n + 2l - 2l - 1 + 1) \\ &= \tfrac{1}{2}(2k - p + 1), \\ j &= (n + l) - (2l + 1) \\ &= k - p, \end{aligned}$$

and substituting for v in equation (22) from equation (18), the resulting differential equation has the form

$$\frac{xd^2g}{dx^2} + (p + 1 - x)\frac{dg}{dx} + (k - p)\, g\,(x) = 0. \tag{23}$$

This equation is satisfied by the *associated Laguerre polynomial of degree* (k − p) *and order* p, (p $\leqslant$ k), which is designated by $L_k^p(x)$. It is analogous to the associated Legendre polynomial of degree k and order m, $P_k^m(x)$, and is derived in a similar manner from the Laguerre polynomial of *zero order* defined thus:

$$L_k^0(x) = \epsilon^x \frac{d^k}{dx^k}(x^k\epsilon^{-x}). \tag{24}$$

This function, it will be observed, bears a distinct similarity to the Hermitian polynomial defined in equation (5.18). It is a polynomial

of degree k in x which is given by the following series:

$$L_k(x) = \sum_{s=0}^{k} (-1)^{k-s} \frac{[k(k-1)\ldots(k-s+1)]^2}{s!} x^{k-s}$$

$$= (-1)^k \left\{ x^k - \frac{k^2}{1!} x^{k-1} + \frac{k^2(k-1)^2}{2!} x^{k-2} \ldots + (-1)^k k! \right\}, \tag{25}$$

and satisfies the differential equation,

$$x \frac{d^2 g(x)}{dx^2} + (1-x) \frac{dg(x)}{dx} + kg(x) = 0.$$

From this, by differentiating p times, equation (23) is derived, which is satisfied by the *associated Laguerre function*,

$$L_k^p(x) = \frac{d^p}{dx^p} L_k(x).$$

For $k = n + l$, and $p = 2l + 1$, the associated Laguerre polynomial of degree $n - l - 1$ $(= k - p)$ and of order $2l + 1$ is given by the series

$$L_{n+l}^{2l+1}(x) = \sum_{s=0}^{n-l-1} (-1)^{s+1} \frac{\{(n+l)!\}^2 x^s}{(n-l-1-s)!\,(2l+1+s)!\,s!}$$

$$= \left\{(n+l)!\right\}^2 \left\{ -\frac{1}{(n-l-1)!(2l+1)!} + \frac{x}{(n-l-2)!(2l+2)!1!} - \frac{x^2}{(n-l-3)!(2l+3)!2!} + \ldots (-1)^{n-l} \frac{x^{n-l-1}}{(n+l)!(n-l-1)!} \right\}. \tag{26}$$

The first ten of these polynomials are as follows:[7]

$$L_0^0(x) = 1;$$

$$L_1^0(x) = -x + 1; \qquad L_1^1(x) = -1$$

$$L_2^0(x) = x^2 - 4x + 2; \qquad L_2^1(x) = 2x - 4, \qquad L_2^2(x) = 2$$

$$L_3^0(x) = -x^3 + 9x^2 - 18x + 6; \qquad L_3^1(x) = -3x^2 + 18x - 18, \qquad L_3^2(x) = -6x + 18, \qquad L_3^3(x) = -6.$$

[7] Condon and Morse, "Quantum Mechanics," p. 63.

Hence the solution of equation (17) is

$$v_{nl} = \epsilon^{-\frac{x^2}{2}} L_{n+l}^{2l+1}(x),$$

and consequently the solution of the S. radial equation (15) is

$$S_{nl}(x) = x^l v_{nl} = \epsilon^{-\frac{x^2}{2}} x^l L_{n+l}^{2l+1}(x), \tag{27}$$

where, according to equations (9), (13), and (14),

$$x = \frac{r}{a_n} = 2r\sqrt{-A_n}$$

$$= 2r \cdot 2\pi \frac{\sqrt{-2\mu E_n}}{h}, \tag{28}$$

and n and l are the two integers designating the eigenfunctions $S(x)$.

It is most convenient to measure r in terms of the radius of the Bohr orbit in the normal state of the hydrogen atom. Introducing the expression of E_n, given in equation (21), and the value of a_0, defined by (3c), into equation (28), the result is

$$x = \frac{2Zr}{na_0}, \tag{29}$$

and in terms of a_0, the eigenvalues are given by

$$E_{n,l} = -\frac{Z^2e^2}{2n^2a_0}. \tag{30}$$

The associated Laguerre polynomials form an orthogonal system, and in the present case it may be shown that the complete expression which must be normalized is of the form [8]

$$x^{l+1}\epsilon^{-\frac{x^2}{2}} L_{n+l}^{2l+1}(x),$$

and that the normalizing factor is given by the relation

$$(N_{n+l}^{2l+1})^2 = \int_0^\infty x^{2l+2}\epsilon^{-x} [L_{n+l}^{2l+1}(x)]^2 \, dx = \frac{2n[(n+l)!]^3}{(n-l-1)!}. \tag{31}$$

If now, in accordance with equation (29), x is replaced by $2Zr/(na_0)$, the *normalized radial function* becomes

$$S_{nl}(r) = \left\{\frac{(n-l-1)!}{2n[(n+l)!]^3}\left(\frac{2Z}{na_0}\right)^3\right\}^{\frac{1}{2}} \left(\frac{2Zr}{na_0}\right)^l \epsilon^{-\frac{Zr}{na_0}} \cdot L_{n+l}^{2l+1}\left(\frac{2Zr}{na_0}\right) \cdots \tag{32}$$

[8] This is due to the fact, as shown in a subsequent section, that the function $S_{nl}^2(r) \cdot r^2$ is required for the physical interpretation.

Table 1 gives the expressions for the normalized functions $S_{nl}(\rho)$, where $\rho = 2Zr/(na_0)$ is used as radial coördinate. The spectroscopic designation of the corresponding energy "term" is given in the first column.

TABLE 1

NORMALIZED RADIAL FUNCTIONS FOR THE ELECTRON IN HYDROGEN-LIKE ATOMS[9]

$$\rho = \frac{2Zr}{na_0}$$

Spectroscopic Notation	n	l	L_{n+l}^{2l+1}	$S_{nl}(\rho)$
$1s$	1	0	-1	$-2\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} \epsilon^{-\frac{\rho}{2}}$
$2s$	2	0	$2\rho - 4$	$\frac{1}{2^{\frac{5}{2}}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} \epsilon^{-\frac{\rho}{2}} (2\rho - 4)$
$2p$	2	1	$-3!$	$-\frac{1}{2\sqrt{6}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} \epsilon^{-\frac{\rho}{2}} \rho$
$3s$	3	0	$-3(\rho^2 - 6\rho + 6)$	$-\frac{1}{3^{\frac{5}{2}}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} \epsilon^{-\frac{\rho}{2}} (\rho^2 - 6\rho + 6)$
$3p$	3	1	$24(\rho - 4)$	$\frac{1}{9\sqrt{6}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} \epsilon^{-\frac{\rho}{2}} \rho(\rho - 4)$
$3d$	3	2	$-5!$	$-\frac{1}{9\sqrt{30}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} \epsilon^{-\frac{\rho}{2}} \rho^2$

Figure 37 shows plots[10] of the eigenfunctions $S_{nl}(r)$ associated with the values $n = 1, 2$, and 3 for the hydrogen atom ($Z = 1$), as functions of r/a_0.

[9] H. Bethe ("Handbuch der Physik," XXIV, Part I, p. 274) and Pauling and Wilson ("Quantum Mechanics," Chapter V) give the expressions for the eigenfunctions corresponding to values of $n > 3$, in which, in order to make the functions positive for small values of r, the sign has been changed. Consequently, the function designated by these writers as $R_{nl}(\rho)$ is identical with $-S_{nl}(\rho)$ as defined by equation (32). Obviously this difference in sign does not affect the value of the distribution function $\{S_{nl}(r) \cdot r\}^2$.

[10] This figure is taken from the treatise by H. E. White, "Introduction to Atomic Spectra." Plots of these functions were first published by L. A. Pauling, *Proc. Roy. Soc.*, **A114**, 181 (1927). See also the summary by the same author in *Chem. Rev.*, **5**, 173 (1928), and L. Pauling and E. B. Wilson, Jr., "Introduction to Quantum Mechanics," Chapter V.

It will be observed that the number of zero points, or nodes (excluding that at $r = 0$), along the axis of r (or ρ) is identical with the value $j = n - (l + 1)$. Hence, j is designated as the *radial quantum number* and

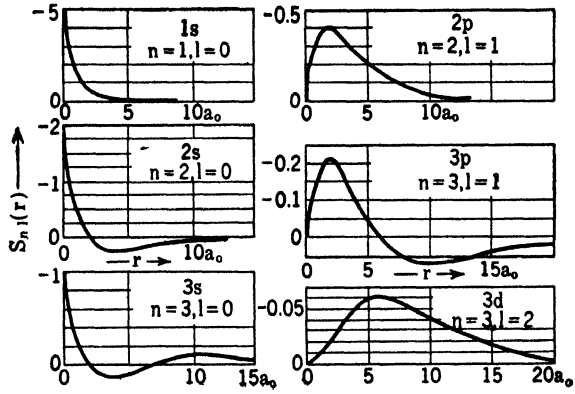


FIG. 37. Plots of normalized radial functions for different electronic states of hydrogen atom.

is often denoted by the symbol n_r. The comments of A. Sommerfeld upon this result are significant.[11] He writes:

> This leads to a simple wave mechanical interpretation of the quantum numbers; indeed, not only in the case of the radial coördinate and of the Kepler problem but in all cases where, using the polynomial method, we can apply the following definition by forcibly breaking off an expansion in series, that is, by the degree of the resulting polynomial: *Quantum numbers denote the numbers of nodes in the proper functions that lie between the limiting points for the coördinate in question.* This brings to mind the analogy of the vibrating string in which the ordinal number of an overtone is likewise measured by the number of nodes that lie between the fixed ends of the string.

7.5 The Complete Solution of the S. Equation for the Hydrogen Atom. As stated at the beginning of this chapter, the complete solution of the S. equation for the hydrogen-like atom must be an expression of the form

$$\phi_{nlm}(r, \theta, \eta) = S_{nl}(r) \cdot Y_{lm}(\theta, \eta),$$

where

$$Y_{lm}(\theta, \eta) = X_{lm}(\theta) \cdot Z_m(\eta)$$

is a tesseral harmonic, which was defined in equation (6.43), with the normalizing factor given by equation (6.46). Furthermore, according to equations (6.19), (6.39), and (6.42), the individual *nor-*

[11] Sommerfeld, "Wave Mechanics," English translation, p. 72.

malized functions of θ and η are given by the relations

$$X_{lm}(\theta) = \left\{\frac{(2l+1)\,(l-m)!}{2(l+m)!}\right\}^{\frac{1}{2}} P_l^m(\cos\theta), \tag{33}$$

$$Z_m(\eta) = \frac{1}{\sqrt{2\pi}} \cdot \epsilon^{im\eta}. \tag{34}$$

Table 2 gives the expressions for these normalized functions for values of $l = 0$, 1, and 2.

TABLE 2

NORMALIZED SPHERICAL EIGENFUNCTIONS

l	m	$X_{l,m}(\theta)$	$Z_m(\eta)$
0	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2\pi}}$
1	0	$\left(\frac{3}{2}\right)^{\frac{1}{2}} \cos\theta$	$\frac{1}{\sqrt{2\pi}}$
1	± 1	$\left(\frac{3}{4}\right)^{\frac{1}{2}} \sin\theta$	$\frac{\epsilon^{\pm i\eta}}{\sqrt{2\pi}}$
2	0	$\frac{1}{2}\sqrt{\frac{5}{2}}\,(3\cos^2\theta - 1)$	$\frac{1}{\sqrt{2\pi}}$
2	± 1	$\frac{3}{2}\sqrt{\frac{5}{3}} \sin\theta\cos\theta$	$\frac{\epsilon^{\pm i\eta}}{\sqrt{2\pi}}$
2	± 2	$\frac{3}{4}\sqrt{\frac{5}{3}} \sin^2\theta$	$\frac{\epsilon^{\pm 2i\eta}}{\sqrt{2\pi}}$

The function $Z_m(\eta)$ can be expressed in either the complex or real form thus:

$$Z_1(\eta) = \frac{1}{\sqrt{2\pi}}\epsilon^{i\eta} \quad \text{or} \quad Z_{1\cos}(\eta) = \frac{1}{\sqrt{\pi}}\cos\eta$$

$$Z_{-1}(\eta) = \frac{1}{\sqrt{2\pi}}\epsilon^{-i\eta} \quad \text{or} \quad Z_{1\sin}(\eta) = \frac{1}{\sqrt{\pi}}\sin\eta$$

$$Z_2(\eta) = \frac{1}{\sqrt{2\pi}}\epsilon^{2i\eta} \quad \text{or} \quad Z_{2\cos}(\eta) = \frac{1}{\sqrt{\pi}}\cos 2\eta$$

$$Z_{-2}(\eta) = \frac{1}{\sqrt{2\pi}}\epsilon^{-2i\eta} \quad \text{or} \quad Z_{2\sin}(\eta) = \frac{1}{\sqrt{\pi}}\sin 2\eta$$

and so forth.

The change in the coefficient from $1/\sqrt{2\pi}$ to $1/\sqrt{\pi}$ takes care of the normalization. By means of the expressions in Tables 1 and 2, it is possible to calculate the expression for the normalized eigenfunction ϕ_{nlm}, corresponding to any given values of n, l, and m. The functions for a few of the lowest states are as follows:

For $n = 1, l = 0, m = 0$, which is the *normal* state for the hydrogen-like atom,

$$\phi_{100} = -\frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}} \epsilon^{-\frac{Zr}{a_0}}. \tag{35}$$

For $n = 2$, there are *four* eigenfunctions, one corresponding to $l = 0$, and three to $l = 1$, as indicated in Table 2. The normalized functions are as follows:[12]

$$\phi_{200} = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{3}{2}}\left(\frac{Zr}{a_0} - 2\right) \cdot \epsilon^{-\frac{Zr}{2a_0}} \tag{36}$$

$$\phi_{210} = -\frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} r\cos\theta \cdot \epsilon^{-\frac{Zr}{2a_0}} \tag{37}$$

$$\phi_{211} = -\frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{\frac{5}{2}} r\sin\theta \cdot \epsilon^{-\frac{Zr}{2a_0}}\epsilon^{\pm i\eta}. \tag{38}$$

From equation (20) it will be observed that *all the eigenfunctions for any given value of* n *have the same eigenvalue*

$$E_n = -\frac{Z^2e^2}{2n^2a_0}.$$

The state is, therefore, of the *degenerate type.* As was shown previously, there are $2l + 1$ functions corresponding to any given value of l. But for a given value of n, l can have the values $n - 1$, $n - 2$, . . . 0. Hence the total number of eigenfunctions corresponding to a given value of n is

$$\Sigma = \{2(n-1)+1\} + \{2(n-2)+1\} + \ldots + 3 + 1 = n^2.$$

Thus the *degree of degeneracy for quantum number* n *is* n^2. This conclusion may be illustrated by the following table giving the different possible values of the quantum numbers n, l, and m, and the corre-

[12] A table of the hydrogen-like wave functions for the values $n = 3$ is given in Pauling and Wilson, "Introduction to Quantum Mechanics," pp. 138–9, also in Ruark and Urey, "Atoms, Molecules and Quanta," p. 564, and Sommerfeld, "Wave Mechanics," p. 71.

sponding spectroscopic designations for the lower energy states of a hydrogen-like atom.

Spectral Designation	*Quantum Number Designation of* ϕ	n	l	m
$1s$	100	1	0	0
$2s$	200	2	0	0
$2p$	210	2	1	0
	211	2	1	± 1
$3s$	300	3	0	0
$3p$	310	3	1	0
	311	3	1	± 1
$3d$	320	3	2	0
	321	3	2	± 1
	322	3	2	± 2

The integer n is known as the *total quantum number*. The number l, which is associated with the azimuthal angle θ, is designated the *azimuthal quantum number*, while m, which is associated with the angle η, is designated the *magnetic* quantum number.

7.6 Quantum Mechanics Interpretation of the Characteristic Functions. In quantum mechanics the eigenfunction ϕ, obtained by solving a S. equation, has no direct physical interpretation. But $\bar{\phi}\phi$ or ϕ^2 (if the function is real) is interpreted as representing a probability of occurrence. In the case of the hydrogen atom, $\phi_{nlm}(r, \theta, \eta)$ is a function of three coördinates. Omitting the subscripts and considering only the normalized functions, the expression

$$\bar{\phi}\phi dv = S^2(r) \cdot X^2(\theta) \cdot Z(\eta)\bar{Z}(\eta)dv$$

is regarded as denoting the probability of occurrence of the electron in the element of volume dv, at the point whose coördinates are r, θ, and η. As shown in Chapter IV,

$$dv = r^2 \sin \theta dr d\theta d\eta,$$

while the element of area on the surface of the sphere at r is

$$dA = r^2 \sin \theta d\theta d\eta.$$

Since $Z(\eta)\bar{Z}(\eta) = 1/(2\pi)$, it is evident that the probability of occurrence of the electron is independent of the angle η.

Hence,

$$P = \frac{X^2}{2\pi}(\theta) = \frac{1}{2\pi N^2}\left\{P_l^m(\cos\theta)\right\}^2$$

is the probability of occurrence of the electron per unit solid angle at any given value of θ. It is evident that this interpretation is equivalent to the statement in Chapter VI that P is a measure of the probability of occurrence per unit area on the surface of a sphere described about the origin of coördinates. That is, the magnitude of P at any given value of θ is a measure of the electron density on the corresponding zone (since P is independent of η).

In Figs. 34 and 35 plots were shown of the function P in the cases $l = 3$, $m = 0$ and $l = 3$, $m = 3$, respectively. Figure 38, taken from

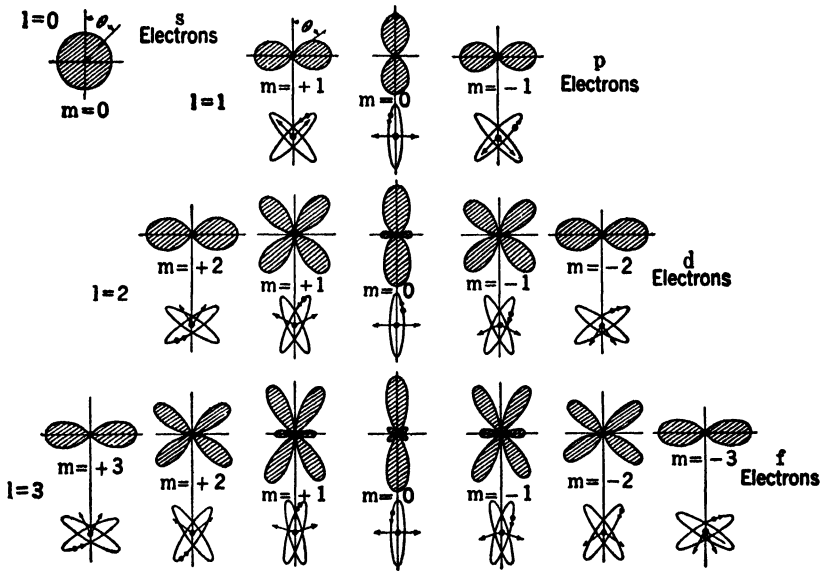


FIG. 38. Angular distribution functions for different electronic states of hydrogen atoms.

the treatise by H. E. White,[13] shows similar plots of the probability density distribution as a function of θ for different electronic states of the hydrogen atom. In the case $n = 1$, $l = 0$, the function P is spherically symmetrical, that is, P is independent of θ. But in the case $n = 2$, $l = 1$, three states are possible, corresponding to $m = 0, \pm 1$. As White describes the plots:

> For these three states, P gives the charge distributions shown at the right and top in Fig. 38. Each curve is shown plotted symmetrically on each side of the vertical axis in order to represent a cross section of the three-dimensional plot. Three-

[13] "Introduction to Atomic Spectra," p. 63. In this book the symbol θ_{ml} is used for the function $P_l^m(\cos\theta)$, and the latter is used to designate the derivative $d^m P_l(x)/dx^m$.

dimensional curves are obtained by rotating each figure about its vertical axis. It should be pointed out that the electron is not confined to the shaded areas in each figure. The magnitude of a straight line joining the center and any point on a given curve is a measure of the electron's probability of being found in the direction of that line.

These figures indicate that for all $m = 0$ states, with the exception of s-electrons, the charge density is greatest in the direction of the poles, i.e., in the direction $\theta = 0$ and π. The exponent of $\epsilon^{im\eta}$ being zero implies that there is no motion in the η-coördinate and that the motion of the electron, i.e., the plane of the orbit, is in some one meridian plane through the η-axis, all meridian planes being equally probable.

For $m = \pm l$, P has its maximum in the direction of the equatorial plane, while for $0 < | m | < l$, P has maxima oriented in definite directions.

These plots correspond very well with the deductions based on the classical Bohr theory for the directions of orientation of the electronic orbits. In a subsequent section it will be shown that, on the basis of the S. theory, the total angular momentum of the electron in its orbit is given by $M = \sqrt{l(l+1)} \cdot h/(2\pi)$ and that the component of angular momentum about the z-axis (the axis through $\theta = 0$ and π) is $M_z = m \cdot h/(2\pi)$. This means that, for absolute values of m greater than zero but less than l, the orbits are oriented with respect to the z-axis.

Figure 38 shows the classical oriented orbits for each state, below the corresponding plot of P. The orbits are drawn slightly tilted out of the normal plane in order to show an orbit rather than a straight line. It should also be added that in the plots " for states $m = 0$ the scale is approximately $1/(l+1)$ times that of the other states having the same value."

Let us now consider the radial function. In this case, the square of the normalized function $\{S_{nl}(r)\}^2$ should evidently be interpreted as the *probability of occurrence of the electron per unit length* at the point whose distance from the nucleus is given by the value of r.

However, a much better physical interpretation of the behavior of the electron in a hydrogen atom is obtained by use of the so-called *distribution function*, designated by D, which is derived from $S(r)$ thus.

The volume of the spherical shell between the radii r and $r + dr$ is evidently given by

$$\frac{d}{dr}\left(\frac{4}{3}\pi r^3\right)dr = 4\pi r^2 dr.$$

Hence,

$$4\pi r^2 \{S_{nl}(r)\}^2 dr = D \cdot dr \qquad (39)$$

is a measure of the *probability of occurrence* of the *electron in a spherical shell of thickness* dr *and of radius* r.

Figure 39, taken from the treatise by H. E. White,[14] gives values of D as functions of r/a_0 for different electronic states of the hydrogen atom. It will be observed that in all cases $D = 0$ for $r = 0$. This

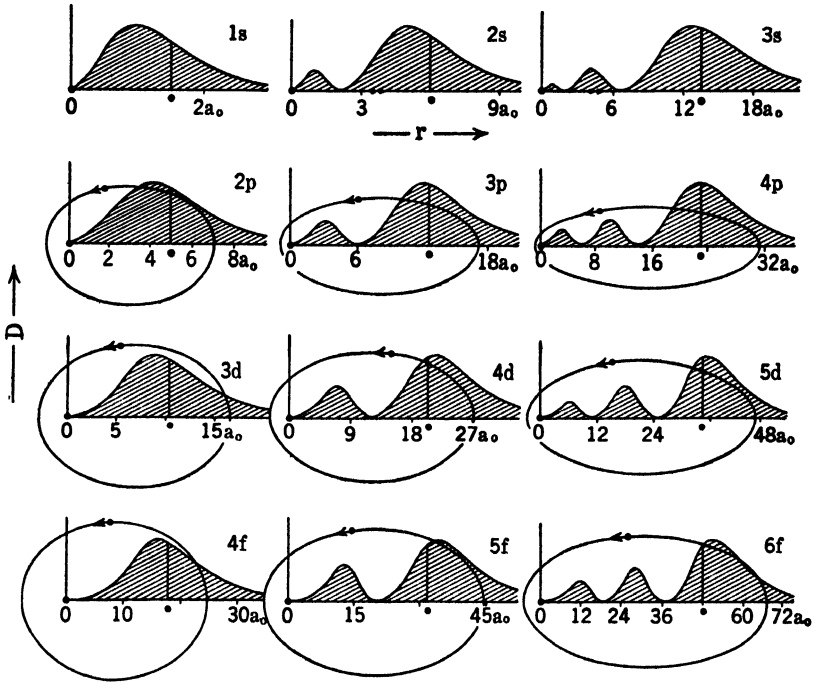


FIG. 39. Radial distribution function (D) for different electronic states of hydrogen atom, and Bohr orbits for comparison.

arises from the fact that the volume of the spherical shell of thickness dr becomes infinitesimally small as r decreases to zero. Stated in other terms, $D = 0$ at $r = 0$ because of the factor r^2.

On the other hand, $\bar{\phi}\phi$ plotted as a function of r, or $S^2(r)$, may be quite large at $r = 0$. This is illustrated by considering the case $n = 1$, $l = 0$ (the normal state of the hydrogen atom). From equation (35) it follows that

$$\bar{\phi}_{100}\phi_{100} = \frac{1}{\pi}\left(\frac{Z}{a_0}\right)^3 \cdot \epsilon^{-\frac{2Zr}{a_0}}.$$

That is, the probability density is a maximum at the origin and decreases

[14] White, *op. cit.*, p. 68. Similar plots are given in Pauling and Wilson, "Quantum Mechanics," p. 143.

exponentially with increase in r. Thus for $r = a_0/2$, the value of $\bar{\phi}\phi$ is $1/\epsilon$ times (36.8 per cent) its value at $r = 0$.

The function D for this state is given by

$$D = 4\pi r^2 \bar{\phi}\phi = 4\left(\frac{Z}{a_0}\right)^3 r^2 \epsilon^{-\frac{2Zr}{a_0}}.$$

It has a maximum at a value $r = r_m$ which may be calculated as follows:

For $$r = r_m, \frac{dD}{dr} = 0.$$

Hence, $$2r_m - \frac{2Zr_m^2}{a_0} = 0.$$

For $$Z = 1, r_m = a_0.$$

Figure 40[15] shows the function ϕ_{100} for the hydrogen atom, at the top, $\rho = \phi_{100}^2$ in the center, and D at the bottom, each plotted as a function of r (in Ångström). The distribution function exhibits a maximum at exactly that value of r which was deduced by Bohr as the radius of the circular electronic orbit in the normal state. On the basis of the Bohr theory, it was also shown that the orbits for the $2p$ and $3d$ states should be circular and of radii $4a_0$ and $9a_0$, respectively. A simple calculation shows that these are exactly the values at which the corresponding D-functions exhibit single maxima, as indicated by the plots in Fig. 39.

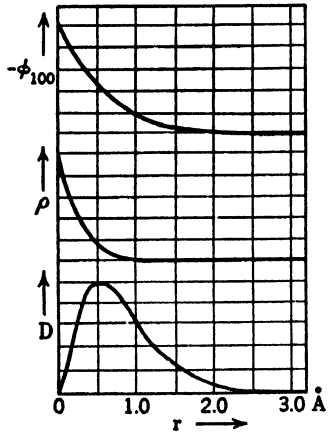


FIG. 40. Plots of characteristic function (ϕ_{100}), density per unit length (ρ), and distribution function (D) for normal state of hydrogen atom.

Figure 41 gives a more graphical interpretation of this function for the case $n = 2$, $l = 0$ (spectroscopic $2s$ state). "If we could imagine the electron in a hydrogen atom replaced by an infinitesimal source of light, the net effect of the fluctuation in the instantaneous location of the electron over a period of time would result in an image which would be brightest at those points where the probability of occurrence is greatest."[16]

[15] Pauling and Goudsmit, "The Structure of Line Spectra," p. 30.

[16] S. Dushman, *J. Chem. Education*, **8**, 1074 (1931).

Figure 42[17] gives a photograph of a three-dimensional representation of the distribution function for $n = 2$, $l = 1$, $m = 0$ (spectroscopic $2p$ state). Most of the charge lies in the region indicated by the boundaries, but actually the density decreases exponentially with r in the space outside the boundaries.

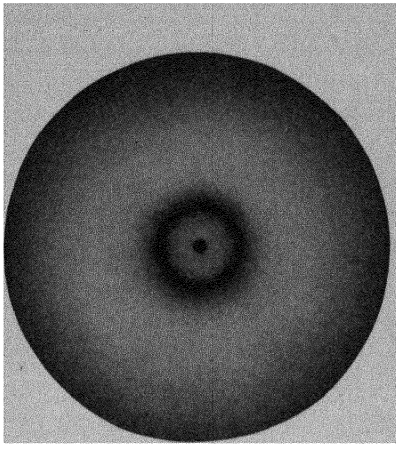


FIG. 41. Illustration of probability of occurrence of electron for excited state of hydrogen atom ($n = 2$, $l = 0$).

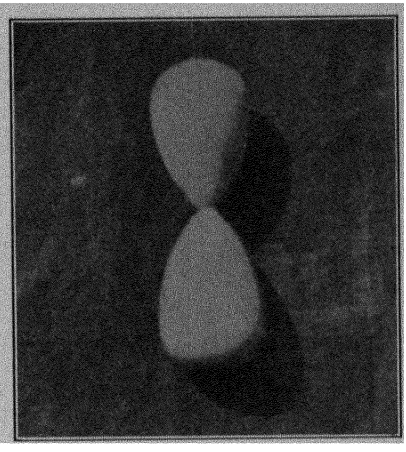


FIG. 42. The charge density distribution for $n = 2$, $l = 1$, $m = 0$. The boundaries of the figure are too sharp, for the charge extends throughout space, though most of the charge lies in the region indicated by the figure.

7.7 Comparison between Deductions from Classical and Quantum Mechanics. The difference between the wave mechanics point of view and that of the Bohr theory is brought out by comparing this conception of a probability density distribution for the electron with that of an electron revolving about the nucleus in a definite elliptic orbit. According to the older theory, which was discussed in Chapter IV, the magnitude of the major axis of ellipse is determined by the value of n, while the ratio of minor axis to major axis is given by the value of k/n where k is designated the azimuthal quantum number. In the new theory, the latter has to be replaced by $\sqrt{l(l+1)}$. With this modification, the corresponding orbits, as deduced from the older theory, are indicated on each of the plots of D in Fig. 39. The origin is taken as one of the foci of each ellipse, so that the distances from the origin along the axis of r to these curves give the maximum and minimum dis-

[17] H. C. Urey, *J. Chem. Education*, **8**, 1114 (1931).

tances of the electron from the nucleus according to the Bohr-Sommerfeld theory.

On the basis of classical mechanics the electron was confined in its motion to an orbit of definite dimensions, determined by the magnitudes of E_n and of k/n. Although the new theory replaces this conception by that of a distribution function for the occurrence of the electron as a function of r, there is this point of resemblance between the two theories, *that the function* D *always exhibits an exponential decrease with increase in* r *beyond the maximum value of the latter which is given by classical theory.*

This conclusion is readily deduced by means of the following argument.

Let us consider the S. radial equation (8) which, when multiplied through by r, becomes

$$r \cdot \frac{d^2S}{dr^2} + \frac{2dS}{dr} + \left\{A + \frac{2B}{r} - \frac{l(l+1)}{r^2}\right\} Sr = 0. \tag{40}$$

Put $R = rS$, so that $R^2 = r^2S^2 = D/4\pi$. Hence,

$$\frac{dR}{dr} = r \cdot \frac{dS}{dr} + S$$

$$\frac{d^2R}{dr^2} = r \cdot \frac{d^2S}{dr^2} + \frac{2dS}{dr},$$

and equation (40) becomes

$$\frac{d^2R}{dr^2} + \left\{A + \frac{2B}{r} - \frac{l(l+1)}{r^2}\right\} R = 0.$$

Substituting for A and B from (9), the last equation becomes

$$\frac{d^2R}{dr^2} + \left\{\frac{8\pi^2\mu}{h^2}\left(E + \frac{Ze^2}{r}\right) - \frac{l(l+1)}{r^2}\right\} R = 0. \tag{41a}$$

From equations (1) and (3c) it is readily seen that

$$\frac{8\pi^2\mu E}{h^2} = -\left(\frac{Z}{na_0}\right)^2; \quad \text{and} \quad \frac{8\pi^2\mu Ze^2}{r} = \frac{2Z}{a_0 r}.$$

Then equation (41a) becomes

$$\frac{d^2R}{dr^2} - \left\{\left(\frac{Z}{na_0}\right)^2 - \frac{2Z}{a_0 r} + \frac{l(l+1)}{r^2}\right\} R = 0. \tag{41b}$$

For very large values of r, the terms containing $1/r$ and $1/r^2$ become negligible, and the differential equation assumes the form

$$\frac{d^2R}{dr^2} - \left(\frac{Z}{na_0}\right)^2 R = 0.$$

The solution of this equation which vanishes for $r \to \infty$ is

$$R = C\epsilon^{-\frac{Zr}{na_0}},$$

which shows that $S = Rr^{-1}$ must decrease exponentially with increase in r.

The condition that the function R shall exhibit a point of inflection is given by $d^2R/dr^2 = 0$. Hence, either $R = 0$ (which is physically untenable) or

$$\frac{l(l+1)}{r^2} - \frac{2Z}{a_0 r} + \left(\frac{Z}{na_0}\right)^2 = 0.$$

Solving this quadratic equation in $(1/r)$ we obtain the two roots

$$\frac{1}{r} = \frac{Z}{l(l+1)a_0}\left\{1 \pm \sqrt{1 - \frac{l(l+1)}{n^2}}\right\}.$$

Hence,

$$r = \frac{na_0}{Z} \cdot \frac{l(l+1)}{n \pm \sqrt{n^2 - l(l+1)}}$$

$$= \frac{na_0}{Z}\{n \pm n^2 - l(l+1)\}. \tag{42}$$

For $l = 0$,

$$r = \frac{2n^2a_0}{Z} \quad \text{or} \quad r = 0.$$

For $l(l+1) = k^2$,

$$r_m = \frac{n^2a_0}{Z}\left(1 + \sqrt{1 - \frac{k^2}{n^2}}\right),$$

$$r_0 = \frac{n^2a_0}{Z}\left(1 - \sqrt{1 - \frac{k^2}{n^2}}\right),$$

where r_m and r_0 are the maximum and minimum values, respectively,

of r. These values correspond to those determined by the Bohr-Sommerfeld theory for orbits of total quantum number n and azimuthal quantum number k.

It is thus evident that the points of inflection on the distribution curves must occur at those values of r which represent the maximum values for the classical orbits. That is, the maximum value of the function D always occurs in the region inside the boundary deduced from classical theory.

There are some other distinct points of similarity in the conclusions derived by means of both the older and newer theories. Thus, as mentioned previously, for the $1s$, $2p$, and $3d$ states the values of D show single maxima at those values of r which are the radii of the corresponding Bohr circular orbits.

A still more striking similarity in the conclusions deduced from the Bohr-Sommerfeld theory with those derived from the calculation of the function D is obtained by comparing the *average value of* r *as computed by each method.*

According to equation (4.98), the average value of r on the basis of the Bohr theory is given by

$$\bar{r} = \frac{1}{\tau}\int_0^{\tau} r\,dt$$

$$= \frac{a_0 n^2}{Z}\left\{1 + \frac{1}{2}\left[1 - \frac{k^2}{n^2}\right]\right\}. \tag{43}$$

For the quantum mechanics case, in accordance with equation (5.30),

$$\bar{r} = \frac{\int \bar{\phi} r \phi\, d\tau}{\int \phi\phi\, d\tau}$$

$$= \frac{\int_0^{\infty} D^2 r\, dr}{\int_0^{\infty} D^2\, dr}, \tag{44}$$

where D, as defined by equation (39), is the probability of occurrence of the electron in a spherical shell of thickness dr at the distance r from the nucleus.

For the normal state of the hydrogen-like atom ($n = 1$, $l = 0$), if we use *normalized* functions, the denominator in (44) is equal to

unity. Hence,

$$\bar{r} = \int_0^\infty 4\pi r^2 \cdot \frac{1}{\pi a_0^3} \cdot \epsilon^{-\frac{2r}{a_0}} r dr$$

$$= \frac{a_0}{4Z} \int_0^\infty \rho^3 \epsilon^{-\rho} d\rho,$$

where $\rho = 2Zr/a_0$.

The value of the integral in the last equation is given in tables of definite integrals (see Appendix III) as 3! Consequently,

$$\bar{r} = \frac{3}{2} \cdot \frac{a_0}{Z}, \tag{45}$$

which is the same as that derived from equation (43) for $k^2 = 0$.

More generally the average value of any power s of r, for any state of the hydrogen-like atom, is given by the relation

$$\bar{r}^s = \int_0^\infty r^s \cdot 4\pi r^2 S_{nl}^2(r) dr. \tag{46}$$

The method of evaluating the integral in this equation has been discussed by I. Waller,[18] and in the case of $s = 1$, the relation derived in this manner has the form [19]

$$\bar{r} = \frac{a_0 n^2}{Z}\left[1 + \frac{1}{2}\left\{1 - \frac{l(l+1)}{n^2}\right\}\right]. \tag{47}$$

It will be observed that this relation is obtained from equation (43) by substituting $l(l+1)$ for the square of the azimuthal quantum number k which was used in the Bohr theory.

7.8 Relation between Kinetic and Potential Energy Deduced from S. equation.[20] For the classical case, the average potential energy is given, as shown in Chapter IV, by the relation

$$\bar{V} = -\frac{Ze^2}{a_n}, \tag{48}$$

and, since E is a constant, the average kinetic energy must be given by the relation

$$\bar{T} = E - \bar{V} = \frac{Ze^2}{2a_n}. \tag{49}$$

[18] *Z. Physik*, **38**, 635 (1926).

[19] White, "Introduction to Atomic Spectra," p. 69; Pauling and Wilson, "Introduction to Quantum Mechanics," p. 144.

[20] The remarks in this section are based on a discussion of this topic in lecture notes by Dr. F. Seitz.

This relation between the two forms of energy is regarded as a balance between the so-called centripetal force of attraction between electron and nucleus and the centrifugal force due to the orbital motions of the two particles. Classical mechanics does not provide any criterion for the determination of stationary states; this is supplied in the Bohr model by the Wilson-Sommerfeld quantum conditions.

In quantum mechanics, on the other hand, we do not associate position with kinetic energy, and the " picture " which we obtain of the atom is that of the " electron cloud " for which the charge distribution is derived from knowledge of the wave function ϕ_n associated with the particular eigenvalue E_n which characterizes any given state of the system. What inferences can we draw from this concept regarding the average values of potential and kinetic energy?

The average value of V is given by

$$\bar{V} = -Ze^2\left(\frac{1}{r}\right) = -Ze^2\int_0^\infty \bar{\phi}_n r^{-1}\phi_n d\tau. \tag{50}$$

Let us consider the case $n = 1$, $l = 0$, for which ϕ_{100} is given by equation (35).

$$\int_0^\infty \bar{\phi}_{100} r^{-1}\phi_{100} d\tau = 4\left(\frac{Z}{a_0}\right)^3 \int_0^\infty r\epsilon^{-\frac{2Zr}{a_0}} dr$$

$$= \frac{Z}{a_0}\int_0^\infty \epsilon^{-\rho}\rho d\rho = \frac{Z}{a_0}. \tag{51}$$

Hence,

$$\bar{V} = -\frac{Z^2e^2}{a_0} = -\frac{Ze^2}{a_1} = 2E,$$

where a_1 is the radius of the corresponding Bohr orbit. (See equations (3*b*) and (3*c*).)

Similarly it may be shown[21] that for any other value of n

$$\left(\frac{1}{r}\right) = \frac{Z}{a_0 n^2} = \frac{1}{a_1 n^2} = \frac{1}{a_n},$$

and

$$\bar{V} = -\frac{Ze^2}{a_n} = 2E,$$

while

$$\bar{T} = \frac{Ze^2}{2a_n} = -E.$$

[21] H. Bethe, "Handbuch der Physik," Vol. XXIV, Part I, pp. 282–286.

That is, the average value of $1/r$ and the average values of the potential and kinetic energies, as derived on the basis of quantum mechanics, are identical with those derived in the Bohr theory.

In ordinary electrostatic theory, the potential energy in the field of a positive charge Ze (at the origin), due to a negative charge distribution defined by the function $e\sigma(r, \theta, \eta)$ is given by

$$V = -Ze^2 \int \frac{\sigma}{r} d\tau.$$

It is evident that equation (50) merely expresses the same relation, in which σ is replaced by $| \phi_n |^2$.

It is also of interest to consider these conclusions, regarding the eigenfunctions and eigenvalues for the hydrogen-like atom, from another point of view which will be found of great assistance in dealing with other atomic and molecular systems.

We can write the S. equation for any system in the form

$$H\phi = \left(-\frac{1}{\alpha^2} \nabla^2 + V \right) \phi = E\phi, \tag{52}$$

where H is designated the *Hamiltonian operator*, and $\alpha^2 = 8\pi^2\mu/h^2$. That is, if ϕ is an eigenfunction of the system, then the result of operating on it with H should be equal to $E\phi$ where E is a constant which is known as the total energy.

Multiplying both sides of (52) by $\bar{\phi}$ and integrating over the configuration space, we obtain the relation

$$\int \bar{\phi} H \phi d\tau = -\frac{1}{\alpha^2} \int \bar{\phi} \nabla^2 \phi d\tau + \int \bar{\phi} V \phi d\tau$$

$$= \int \bar{\phi} E \phi d\tau.$$

Since V is a function of the coördinate variables, E is a constant, and ϕ is assumed to be normalized, the last equation can be written in the form

$$-\frac{1}{\alpha^2} \int \bar{\phi} \nabla^2 \phi d\tau + \int V \bar{\phi} \phi d\tau = E. \tag{53}$$

As shown already, the second integral corresponds to the average potential energy, and therefore the first integral must correspond to the average kinetic energy of the particles constituting the system. In a subsequent chapter it will be shown that if the form of the function ϕ

is chosen in such a manner as to make E a *minimum*, then ϕ will be a solution of the S. equation (52).

Now let us consider the factors which govern the actual values of each of the integrals in equation (53). For the purpose of such discussion it is convenient to assume that ϕ is a real function of the three rectangular coördinates. However, the conclusions which we shall derive are equally applicable to the case in which ϕ is a complex function of any other set of curvilinear coördinates.

In terms of rectangular coördinates, the left-hand integral in equation (53), which we shall designate by I_T, has the form

$$I_T = -\frac{1}{\alpha^2}\int \phi \left\{\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right\}\phi d\tau.$$

Now

$$\frac{d}{dx}\left(\frac{\phi\partial\phi}{\partial x}\right) = \left(\frac{\partial\phi}{\partial x}\right)^2 + \frac{\phi\partial^2\phi}{\partial x^2}.$$

Hence

$$\frac{\phi\partial\phi}{\partial x}\Big]_{-\infty}^{\infty} = \int_{-\infty}^{\infty}\left(\frac{\partial\phi}{\partial x}\right)^2 dx + \int_{-\infty}^{\infty}\frac{\phi\partial^2\phi}{\partial x^2}\cdot dx.$$

But in order that ϕ shall be a " sensible " solution of the S. equation, this function must vanish at the limits $x = \pm\infty$, and the same conditions must apply with respect to the y and z coördinates. Hence, we obtain the relation

$$I_T = \frac{1}{\alpha^2}\int\left\{\left(\frac{\partial\phi}{\partial x}\right)^2+\left(\frac{\partial\phi}{\partial y}\right)^2+\left(\frac{\partial\phi}{\partial z}\right)^2\right\}d\tau,$$

which shows that I_T is *positive*. Now $\partial\phi/\partial x$, $\partial\phi/\partial y$ and $\partial\phi/\partial z$ correspond to the components of the gradient of the function ϕ with respect to each of the axes of coördinates, where we define the *gradient* $d\phi/dn$, as the rate of change in ϕ with distance measured along the normal to the surface representing the function ϕ. (See Appendix IV.)

Hence, the last relation can be written in the form

$$\bar{T} = I_T = \frac{1}{\alpha^2}\int\left(\frac{d\phi}{dn}\right)^2 d\tau. \tag{54}$$

It is evident that in the case of a spherically symmetrical function, such as that for the lowest state of the hydrogen atom, $d\phi/dn$ will be identical with $d\phi/dr$.

Since I_T is positive and the term involving V in equation (53) is

negative for stable states of the system, it is evident that these two terms tend to counterbalance each other. In order to minimize E we might attempt to minimize $\bar{V}$ by locating the electron in the hydrogen-like atom as near the nucleus as possible, since this would increase the absolute value of Ze^2/r. But this will tend to make $\bar{T}$ very large because of the increased value of $d\phi/dn$. That is, the function ϕ will change so rapidly with decrease in r that $d\phi/dn$ will assume very large values and E will not be as negative as possible.

On the other hand, if the form of ϕ is chosen in such a manner as to make $\bar{T}$ small, this means small values of $d\phi/dn$ and a consequent "spread" over a large range of values of r in the function ϕ and in the corresponding charge distribution function $e\bar{\phi}\phi$. Such a function will not, however, give a charge distribution which is localized in the most satisfactory manner in the regions of negative V, with the result that, again, E will not be as low as possible. The best wave function will therefore have a form which is intermediate between these two extremes, and as shown previously, the compromise actually obtained is such that

$$\bar{T} = -\tfrac{1}{2}\bar{V}, \quad \text{and} \quad E = \bar{T} + \bar{V} = \tfrac{1}{2}\bar{V}.$$

We can draw the same conclusions regarding the necessity for a compromise between $\bar{T}$ and $\bar{V}$ from the Principle of Indeterminacy. For this purpose we shall write the expression for the average kinetic energy in the form

$$\bar{T} = \text{Av. of } \frac{1}{2\mu}(p_x^2 + p_y^2 + p_z^2),$$

where p_x, p_y, p_z are the components of momentum with respect to the three coördinate axes, and we consider the mean value of the sum of the squares of these components. From the relation

$$\Delta p \cdot \Delta q \gtrsim h$$

it is seen that if we choose ϕ in such a manner as to make Δq very small (that is, confine the electron to a limited region of space, as in the Bohr theory) then Δp must become very large, and since the average value of T for one coördinate is approximately $(\Delta p)^2/(2\mu)$, the kinetic energy will also increase enormously. On the other hand, the potential energy will vary as $1/\Delta q$, that is, $\bar{V}$ will be lower (more negative), the smaller Δq. The best compromise is obtained by choosing such a form for ϕ that the charge will be spread over a fairly extensive range of values of r, so that, as a consequence, the magnitude of either $\bar{T}$ or $\bar{V}$ will not become excessively large.

These considerations will appear as specially important when we

come to deal with atomic systems involving more than one electron and with molecular systems. In these and similar cases, the distribution functions for the electrons must satisfy the condition that the total energy of the system, as derived from an equation such as (53), shall be a minimum.

7.9 Angular Momentum of Electron. In the last section of Chapter VI it was deduced that the average value of M^2 may be computed from the relation

$$\left(M^2 + \frac{h^2}{4\pi^2}\Omega\right)\phi = 0,$$

where Ω is an operator defined in terms of θ and η.

In order to determine whether the electron in state n, l, m possesses a definite value of M, it is necessary to find whether the relation

$$\left(M^2 + \frac{h^2}{4\pi^2}\Omega\right)\phi_{nlm} = 0$$

has solutions of the form

$$M^2\phi_{nlm} = a\phi_{nlm},$$

where a is a constant. In that case the solution is $M^2 = a$.

For the electron in the hydrogen-like atom

$$\phi_{nlm} = S_{nl}(r) \cdot Y_{lm}(\theta, \eta).$$

Hence,

$$M^2\phi_{nlm} = M^2 S_{nl}(r) \cdot Y_{lm}(\theta, \eta)$$

$$= -\frac{h^2}{4\pi^2} \cdot S_{nl}(r) \cdot \Omega Y_{lm}(\theta, \eta),$$

since Ω is an operator which does not involve r.

But from the solution of equation (6.14) it follows that

$$\Omega Y_{lm} = -l(l+1)Y_{lm}.$$

Consequently,

$$M^2\phi_{nlm} = \frac{h^2}{4\pi^2} \cdot l(l+1)Y_{lm}(\theta, \eta)S_{nl}(r)$$

$$= \frac{h^2 l(l+1)}{4\pi^2} \cdot \phi_{nlm}.$$

Hence we conclude that the solution is

$$M^2 = \frac{h^2 l(l+1)}{4\pi^2}. \tag{55}$$

That is, in the state of quantum number l, the total angular momentum is $\sqrt{l(l+1)}$ in units of $h/2\pi$. In a similar manner it may be deduced, as in the case of the rigid rotator, that the angular momentum with respect to the z-axis is

$$M_z = \frac{mh}{2\pi}. \tag{56}$$

COLLATERAL READING

The most complete treatment of the hydrogen atom system is given in Pauling and Wilson's " Quantum Mechanics," Chapter V, and by H. Bethe, in " Handbuch der Physik," XXIV, Part I, pp. 274–289. The latter discusses the Laguerre functions very fully.

In connection with section 9, the readers should consult the discussion by Slater and Frank, "Theoretical Physics," Chapter XXIII, and also the remarks by H. E. White, "Introduction to Atomic Spectra," Chapter IV.

CHAPTER VIII

VAN DER WAALS FORCES

8.1 Van der Waals' Equation. An ideal or "perfect" gas obeys the equation of state

$$PV = nRT, \tag{i}$$

where P is the pressure; V, the volume; n, the number of moles; T, the absolute temperature; and R, the gas "constant." For P in atmospheres and V in cubic centimeters, $R = 82.05$ (cm.^{3} atmos. deg.$^{-1}$ mole^{-1}). This equation is valid for low pressures and comparatively high temperatures, where forces of attraction and repulsion between molecules in the gas have a negligible effect. But at higher pressures or low temperatures, especially near those of condensation to the liquid state and at pressures of the order of an atmosphere or higher, it has been found that equation (i) is not in agreement with the observed data. A very large number of empirical or semi-theoretical equations of state have been suggested, but that postulated by van der Waals is the best known and has been accepted most generally. This equation has the form

$$\left(P + \frac{a}{V_m^2}\right)(V_m - b) = RT, \tag{ii}$$

where V_m is the volume per mole, and a and b are constants whose values are determined from actual observations on the relation between P, V_m, and T over a range of pressures and temperatures.

In this equation the constant b is given by the relation

$$b = 4Nv = \frac{2\pi N r_0^3}{3},$$

where N = Avogadro's number = 6.064×10^{23} molecules per mole;[1] v = volume of a molecule = $(\frac{1}{6})\pi r_0^3$, where r_0 = distance between centers of molecule = *molecular diameter*.

The constant a is derived from the following considerations. The net force acting between two molecules is the resultant of a force of attrac-

[1] Values of physical constants used by the writer are those given by R. T. Birge, *Phys. Rev. Suppl.*, **1**, 1 (1929); also "Smithsonian Physical Tables," edited by F. E. Foule, **1933**, pp. 73–107. Also see Appendix II.

tion which varies with the distance according to a law of the form

$$F_A = -\frac{A}{r^m},$$

and a force of repulsion, which may be written in the form

$$F_R = \frac{B}{r^n},$$

where A, B, m, and n are constants for any given pair of molecules. Hence the total force acting between the two molecules is

$$f = -\frac{A}{r^m} + \frac{B}{r^n}. \tag{iii}$$

The mutual potential energy $U(r)$ is given by

$$U(r) = -\int_{\infty}^{r} f dr = -\frac{A}{(m-1)r^{m-1}} + \frac{B}{(n-1)r^{n-1}}. \tag{iv}$$

If the molecules (or atoms) condense to form a solid lattice, they will assume positions of equilibrium such that $f = 0$, and consequently $dU/dr = 0$, that is, the potential energy will be a *minimum* for the position of equilibrium. Let r_0 be the value of r for this condition. Then

$$\frac{A}{r_0^m} = \frac{B}{r_0^n}, \tag{v}$$

and the mutual potential energy becomes

$$U_0 = -\int_{\infty}^{r_0} f dr = -\frac{A}{r_0^{m-1}} \frac{(n-m)}{(n-1)(m-1)}. \tag{vi}$$

Since U_0 is negative, it follows that n must be greater than m. An illustration of such potential energy plots is given at the end of this chapter (Fig. 51).

In terms of $U(r)$, as determined by means of equation (iv), the constant a is given by the relation[2]

$$a = -\frac{2\pi N^2}{1.013 \times 10^6} \int_{r_0}^{\infty} U(r) \cdot r^2 dr, \tag{vii}$$

where a is in atm. cm.^{6} mole^{-2}, and 1 atmosphere = 1.0132×10^6 dynes/cm.^{2}

[2] F. London, *Z. Physik*, **63**, 245 (1930). See also references to collateral reading at the end of the chapter.

It should be added that according to van der Waals

$$\left.\begin{aligned} b &= \frac{RT_c}{8P_c} \\ a &= \frac{27}{64} \cdot \frac{R^2T_c^2}{P_c} \end{aligned}\right\}, \qquad \text{(viii)}$$

where P_c and T_c are the critical pressure and temperature, respectively.

The problem regarding the origin of these attractive forces between molecules engaged the attention of many investigators and various explanations were suggested. Each of these hypotheses was, however, found to be unsatisfactory, and it was only on the basis of wave mechanics that a satisfactory theory was ultimately developed by F. London[3] in 1930. According to this point of view, the van der Waals forces of attraction are shown to be due to a dynamical polarization of each molecule in the electric field due to its neighbor. The motion of the electrons in one molecule or atom modifies that of the electrons in the other molecule or atom in such a manner that the *electrons tend on the average to move in phase.* The forces that arise are thus similar in nature to those between dipoles. (A dipole is a system consisting of a positive and negative charge located at a definite distance from each other. If e denotes the magnitude of the charge at each " pole," and l, the distance between the point-charges, the dipole moment is given by $\mu_e = el$.) Only in the case of attraction between atoms or molecules the dipole moments fluctuate between maximum and minimum values and the relative orientations vary continuously.

The theory of London and similar considerations by J. C. Slater and J. G. Kirkwood[4] lead to the conclusion that the van der Waals attractive force between two molecules varies inversely as the seventh power of the distance. On this basis it has been found possible to calculate from the electron configurations of monatomic molecules (e.g., the rare gases) values of the constant a which are in satisfactory agreement with the values deduced from the equation of state or from critical data.

In order to understand better the point of view and the arguments developed by London, we shall consider first a phenomenon in classical mechanics which presents a certain similarity, from the mathematical

[3] F. London, *loc. cit.*, also *Z. physik. Chem.*, **B11**, 222 (1931). It should be mentioned that S. C. Wang, *Physik. Z.*, **28**, 663 (1927), first suggested the theory of dipole interaction, but his method of calculation led to results which have not been confirmed by subsequent investigators.

[4] Slater and Kirkwood, *Phys. Rev.*, **37**, 682 (1931).

point of view, to the interaction of two atoms as postulated by the wave mechanics theory.

8.2 Two Interacting Pendulums. Whenever we have two vibrating systems which are coupled, however weakly, it is possible for an interchange of energy to occur between them. The simplest illustration of this phenomenon is the behavior of two *identical* pendulums mounted, adjacent to each other, on a support which permits the motion of one pendulum to be transmitted to the other. The following description of the behavior of the pendulums is taken from the discussion by C. G. Darwin.[5]

We set the pendulum A in motion, while B is at rest. If the support were quite rigid A would go on swinging, and B would stay at rest. But there is a small effect on B through the give of the support, and consequently B begins to move. What happens is rather striking, for B starts swinging more and more and at the same time A's motion diminishes until B is swinging to the full amount while A has come to rest. Afterwards the motion is transferred back to A again, and it continues passing from one to the other until the motion has died away altogether.

The reason for this behavior is easily understood. The system of the two pendulums, like every other vibrating system, has normal modes of vibration, but these do not consist one in the motion of A and the other in the motion of B. The modes are shown in Fig. 43. In one of them A and B are swinging equally both

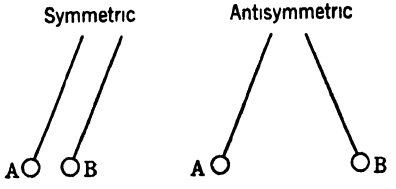


FIG. 43. Illustrating the two modes of vibration of two interacting pendulums.

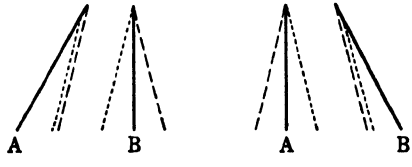


FIG. 44. Motion of each pendulum as a superposition of the two modes.

to the left at the same time, and in the other they are also swinging equally, but now when one is to the left the other is to the right. To distinguish the two modes I shall call them by names of which the application is not very obvious for the pendulums, but which are convenient because they are used in the quantum theory. The first is the symmetric mode, the second the antisymmetric. The two modes have frequencies of vibration which differ from one another a little, because the effect of the yielding support is different in the two cases. The motion we described before in which B and A alternately come to rest is the superposition of these two vibrations. At one time the phases of the two modes are such that A is to the left at the same time for both. Then B would be to the left for one and to the right for the other and so by the principle of superposition it is at rest. Later, on account of the slight difference of the frequencies of the modes, it will occur that the two modes are in opposite phase for A, and therefore in the same phase for B, and consequently the motion will now have passed entirely to B. The exchange of the motions

[5] Darwin, "The New Conceptions of Matter," Chapter VIII.

is illustrated in Fig. 44. It is perhaps well to notice that in likening the exchange of motion of the pendulums to the beats of the two piano wires, the analogy is not between each wire and each pendulum, but between each wire and each mode of the two pendulums. It is the two modes that beat together.

The relation between the frequencies of the two modes and that of each pendulum when not affected by the presence of the other and the character of the motion of each pendulum when interaction occurs are described in the following remarks by N. F. Mott.[6]

The frequencies of the system in these two normal modes are not equal to one another, and neither is equal to the frequency with which either pendulum would oscillate, were the other one absent. Let us call this latter frequency ν and the frequency with which the pendulums vibrate in phase ν_1, and with opposite phase ν_2. One can see that ν_2 is greater than ν and ν_1 less than ν. For when they are swinging in opposite phase, the pendulums are, so to speak, pulling one another back all the time, and so increasing the restoring couple. The frequency must therefore be greater than it would be in the absence of one of the pendulums. When the pendulums are swinging in the same phase, the opposite is the case; the pendulums help each other to swing outwards, the restoring couple is decreased, and so the frequency is less than it would otherwise be. One can also see that the amount by which ν_1 and ν_2 differ from ν depends on the strength of the coupling between the pendulums. For instance, if the string on which the pendulums are hung be fairly tight, and the two pendulums are hung very near to opposite ends of the string, then ν_1 and ν_2 will not differ from ν by as much as they would if the pendulums were hung nearer to the middle of the string.

It is on the existence of these two different frequencies ν_1 and ν_2 that the slow exchange of energy from one pendulum to the other and back depends. If any system is capable of vibrating in a certain number of normal modes, then the most general vibration possible is obtained by superimposing these normal modes, one upon the other. In our case, if θ_1, θ_2 are the angles through which the two pendulums have swung at any moment, then the equations, giving θ_1 and θ_2, for the first normal mode (when the two vibrate in phase) are

$$\theta_1 = A \cos 2\pi(\nu_1 t + \alpha), \quad \theta_2 = A \cos 2\pi(\nu_1 t + \alpha).$$

In the second normal mode (in which the pendulums swing in opposite phase)

$$\theta_1 = B \cos 2\pi(\nu_2 t + \beta), \quad \theta_2 = -B \cos 2\pi(\nu_2 t + \beta).$$

A and B are arbitrary amplitudes, α, β are arbitrary phases. The most general possible vibration, obtained by adding together these two, is

$$\left.\begin{aligned} \theta_1 &= A \cos 2\pi(\nu_1 t + \alpha) + B \cos 2\pi(\nu_2 t + \beta) \\ \theta_2 &= A \cos 2\pi(\nu_1 t + \alpha) - B \cos 2\pi(\nu_2 t + \beta) \end{aligned}\right\}. \qquad \text{(ix)}$$

The angular velocities of both pendulums can be found by differentiating these equations with respect to t. Now suppose we start the system swinging at time $t = 0$, with one pendulum at rest in its position of equilibrium, and the other pulled aside through an angle C. Then at time $t = 0$, say, we must have

$$\theta_1 = \dot{\theta}_1 = 0, \theta_2 = C, \dot{\theta}_2 = 0.$$

[6] Mott, " An Outline of Wave Mechanics," Chapter VI.

The constants A, B, α, β are therefore determined. We find that

$$\alpha = \beta = 0; A = -B = \frac{C}{2}.$$

Writing (ix) in a slightly different form, then, we have for θ_1, at any time t,

$$\theta_1 = -C \sin 2\pi\left(\frac{\nu_1 + \nu_2}{2}\right)t \sin 2\pi\left(\frac{\nu_1 - \nu_2}{2}\right)t, \tag{x}$$

and for θ_2

$$\theta_2 = C \cos 2\pi\left(\frac{\nu_1 + \nu_2}{2}\right)t \cos 2\pi\left(\frac{\nu_1 - \nu_2}{2}\right)t. \tag{xi}$$

Now, ν_1 and ν_2 are only slightly different from ν and from each other. Therefore the second factor in (x) and (xi) has a very much longer period than the first. Each of the expressions (x) and (xi) then represents a vibration with frequency very nearly equal to ν, and with varying amplitude

$$C \begin{matrix}\sin \\ \cos\end{matrix} 2\pi \frac{(\nu_1 - \nu_2)}{2} t.$$

While the amplitude of the one increases, the amplitude of the other decreases. In fact, a measurement of the time that it takes for the energy to go over from one pendulum to the other and to come back again amounts to a measurement of $\nu_1 - \nu_2$. If the coupling between the two pendulums is strong, this time will be shorter than if it were weak.

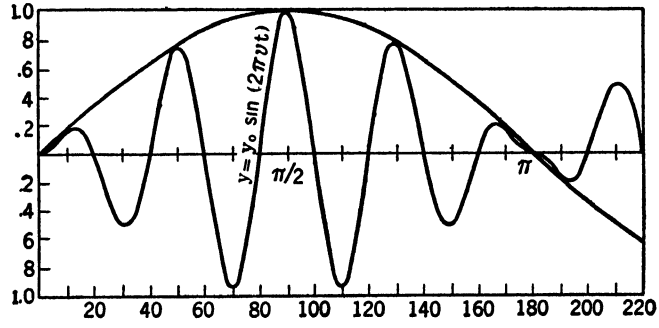


FIG. 45. Plot of the function $y = \sin 2\pi\nu t \cdot \sin 2\pi(\nu t/9)$.

It is interesting to realize the interpretation of the expressions for θ_1 and θ_2. Figure 45 gives an illustration of such a vibration in which

$$\theta_1 = y = \sin 2\pi\nu t \cdot \sin \frac{2\pi\nu t}{m} = y_0 \sin 2\pi\nu t,$$

where $\nu/m = (\nu_1 - \nu_2)/2$, $\nu = (\nu_1 + \nu_2)/2$, and $m = 9$. That is, the period ($\tau = 1/\nu$) of the vibration of frequency $(\nu_1 - \nu_2)$ is nine times that of the frequency $(\nu_1 + \nu_2)$. The expression for θ_2 in equation (iii) gives the same curve, but for $t = 0$, it has a maximum value

corresponding to the abscissa $\pi/2$, instead of the value 0 which is obtained in the case of θ_1 for $t = 0$.

8.3 Electron in Two Adjacent "Boxes." As mentioned in Chapter III, an electron in a potential "box" resembles in its behavior that of the electron in a hydrogen atom. The electron in the box possesses a series of discrete energy states because it is only for these states that stationary de Broglie waves are obtained by the reflection of the wave motion at the walls. For an electron between barriers a distance $2a$ apart, the energy states are given by the relation

$$E_m = \left(\frac{m}{2a}\right)^2 \cdot \frac{h^2}{8\mu}, \tag{8.29}$$

and according to equation (8.30), the eigenfunction for the lowest state E_1 is

$$\phi_1 = 2A \cos \alpha x,$$

which determines the probability amplitude for the electron inside the barriers, while outside these barriers the eigenfunction is

$$\phi_{II,\,III} = 2A \cos \alpha a \cdot \epsilon^{\pm\beta(x-a)}.$$

In these expressions,

$$\alpha^2 = \frac{8\pi^2\mu E}{h^2}; \quad \beta^2 = \frac{8\pi^2\mu(U - E)}{h^2},$$

where $E = E_1$ (in this case) and U ($\gg E$) is the value of the potential energy outside the barriers.

Now let us consider the behavior of an electron in a potential energy field such as that shown in Fig. 46(*a*), which corresponds to two potential boxes (of atomic dimensions) adjacent to each other.[7] Evidently, if the compartments are very far apart (relative to the dimensions of the compartments), the electron will remain for a very long period in the particular region in which it is placed. But as the distance between the boxes is decreased, there will be an increasing probability of penetration from one region into the other, and for relatively small distances (of the order of the de Broglie wave length) there is an equal probability for the occurrence of the electron in either box, if observations are carried out over a sufficiently long period.

This is due to the fact that the ϕ-"pattern" for each energy state extends over all the space outside the barriers, and there will be a portion of the cosine curve in each compartment. In Fig. 46(*b*) the curves

[7] See non-mathematical discussion of this case by R. W. Gurney, "Elementary Quantum Mechanics," pp. 44–48.

MFK and *NGH* show that each ϕ-plot extends into the other box when the two boxes are a distance $2b$ apart, where b is comparable with a; the ordinates of the exponential portions at the edge of the other box (*BK* or *AH*) are no longer vanishingly small. There is a possibility of interaction, analogous to that observed in the case of the two pendu-

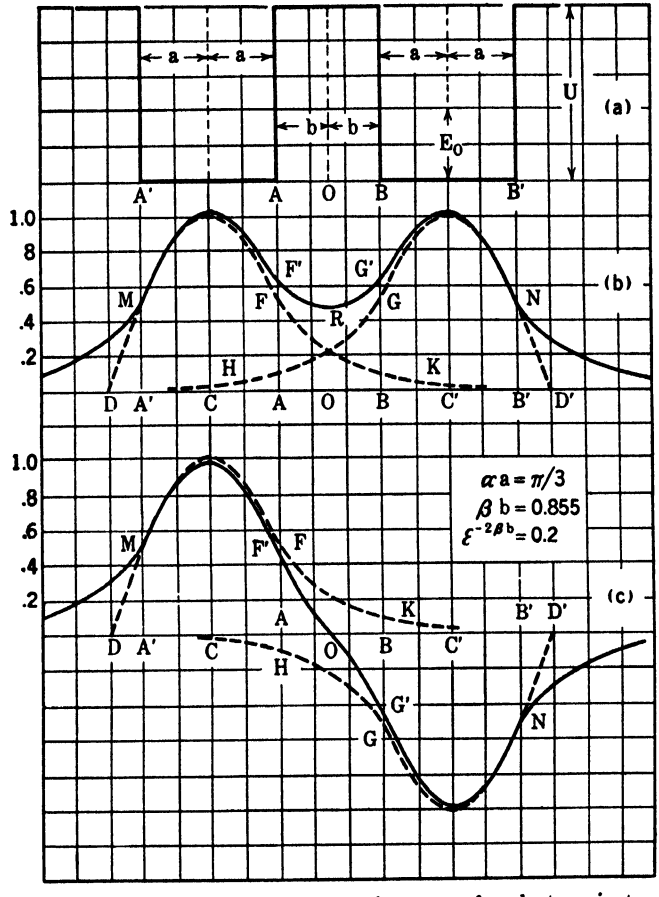


FIG. 46. Illustrating the two eigenfunctions for electron in two adjacent "boxes."

lums, and the two ϕ-curves which were quite separate for the boxes at an infinite distance will merge into one.

When we consider the problem of joining these two curves by a curve which shall be continuous and of finite amplitude at all points between the two boxes, it is seen that there are two solutions possible. These are shown in Fig. 46 (*b*) and (*c*), corresponding respectively to the *symmetrical* and *antisymmetrical modes of vibration* or eigenfunctions.

In the former case, the ϕ-curves for each compartment are joined by a portion which corresponds to the hyperbolic cosine

$$\cosh \beta x = \tfrac{1}{2}(\epsilon^{\beta x} + \epsilon^{-\beta x});$$

in the second, the intermediate portion corresponds to the hyperbolic sine

$$\sinh \beta x = \tfrac{1}{2}(\epsilon^{\beta x} - \epsilon^{-\beta x}).$$

(These functions have been discussed in Chapter II, and illustrated in Fig. 6.)

Now let us consider first the *symmetric* case. The value of ϕ (for the lowest energy state) at A is given, in the case of an electron in a single compartment, by the ordinate $AF = 2A \cos \alpha a$, and the exponential part FK by the relation

$$\phi_{II} = 2A \cos \alpha a \cdot \epsilon^{-\beta(x-a)},$$

where the origin of abscissa is at C. Hence at $x = 2b + a$,

$$\phi_{II} = 2A \cos \alpha a \cdot \epsilon^{-2\beta b}$$

which corresponds to the ordinate BK. Since the ϕ-patterns are symmetrical about O, $BK = AH$. The value of ϕ in the region AOB will be given by the sum of the values which are due to the exponential portions FK and GH.[8] This is indicated by the curve $F'RG'$, so that $F'A = FA + AH$, and hence

$$F'A = G'B = 2A \cos \alpha a(\epsilon^{-2\beta b} + 1).$$

But if the function ϕ and its derivative $d\phi/dx$ are to be continuous over the whole range of values of x, it is evidently necessary that the ordinate of the cosine curve inside the box shall also be increased. That is, the ordinate AF' must correspond to a value $\phi_I = 2A \cos \alpha_1 a$ where α_1 is different from α. In other words, the necessity for continuity of the ϕ-pattern leads to a change in the value of $\alpha^2 = 8\pi^2\mu E/h^2$, and consequently to a *change in the value of* E for the energy state of the electron. We thus derive a relation between α_1 and α, of the form

$$2A \cos \alpha_1 a = 2A \cos \alpha a(1 + \epsilon^{-2\beta b}),$$

that is,

$$\frac{\cos \alpha_1 a}{\cos \alpha a} = 1 + \epsilon^{-2\beta b}. \tag{1}$$

Thus $\cos \alpha_1 a$ is slightly greater than $\cos \alpha a$, and therefore α_1 *must be less than* α. ($\cos 0 = 1$, $\cos \pi/2 = 0$.) That is, the *value* E_s, for the symmetric case, *is less than* E. Equation (1) may be expressed in a more convenient form thus:

[8] This is an illustration of the application of the Principle of Superposition, discussed in Chapter II.

Let $\alpha_1 = \alpha - \Delta\alpha$, where $\Delta\alpha$ denotes a very small change in α. Then,

$$\frac{\cos(\alpha - \Delta\alpha)a}{\cos \alpha a} = \frac{\cos \alpha a \cdot \cos(a\Delta\alpha) + \sin \alpha a \cdot \sin(a\Delta\alpha)}{\cos \alpha a}$$

$$= 1 + \frac{\sin \alpha a}{\cos \alpha a}(a\Delta\alpha) = 1 + \beta a\left(\frac{\Delta\alpha}{\alpha}\right), \tag{2}$$

since $\cos(a\Delta\alpha) = 1$, approximately, and $\sin(a\Delta\alpha) = a\Delta\alpha$, approximately, and according to equation (3.21), $\tan \alpha a = \beta/\alpha$. From equations (1) and (2) it follows that

$$\frac{\Delta\alpha}{\alpha} = \frac{\epsilon^{-2\beta b}}{\beta a}. \tag{3}$$

Since $2\alpha\Delta\alpha = 8\pi^2\mu \cdot \Delta E/h^2$ where ΔE = *decrease in energy*, it is evident that the last equation may be replaced by the relation

$$\frac{\Delta E}{E} = \frac{\lambda \cdot \epsilon^{-\frac{4\pi b}{\lambda}}}{\pi a}, \tag{4}$$

where $$\lambda = \frac{h}{\sqrt{2\mu(U - E)}} = \frac{2\pi}{\beta}$$

represents a distance.

Thus, suppose $b = a$ and $\lambda = 4a = 4b$. Then,

$$\frac{\Delta E}{E} = \frac{4\epsilon^{-\pi}}{\pi} = 0.055.$$

If $2a = 2 \times 10^{-8}$ cm., $E = 9.15$ electron volts, as shown in Chapter III, and $\Delta E = 0.50$ electron volt. Since $\lambda = 12.21 \times 10^{-8}/\sqrt{V}$ cm., where V is the kinetic energy in electron volts,[9] the value $\lambda(= 4a = 4 \times 10^{-8}$ cm.) corresponds to a value for $U - E$ of $(12.21/4)^2 = 9.38$ electron volts.

[9] Since $\lambda = h/\sqrt{2\mu E}$, and $E = Ve$, where V is the potential difference required in order that the electron shall acquire the given kinetic energy, it follows that

$$\lambda = \frac{6.55 \times 10^{-27}}{\left(2 \times 9 \times 10^{-28} \times 4.77 \times 10^{-10}\frac{V}{300}\right)^{\frac{1}{2}}} = \frac{12.21 \times 10^{-8}\text{ cm.}}{\sqrt{V}}$$

where
$e = 4.77 \times 10^{-10}$ electrostatic units;
$\mu = 9 \times 10^{-28}$ g.;
$h = 6.55 \times 10^{-27}$ erg · sec.;
$\frac{1}{300}$ = conversion factor from ordinary volts to electrostatic units of potential.

The *antisymmetric* case is treated quite similarly. The full curve Fig. 46(*c*) shows the hyperbolic sine portion $F'OG'$. In this case, $F'A = FA - FF' = FA - BK$, so that we obtain the relation

$$\frac{\cos \alpha_2 a}{\cos \alpha a} = 1 - \epsilon^{-2\beta b}. \tag{5}$$

Hence α_2 must *be greater* than α. That is, the *value of* E_A, the energy of the electron for the antisymmetric case, *is greater than* E, and as in the symmetric case it is seen that the *increase in energy* ΔE is given, as before, by equation (4).

Thus, it is seen that the original ground level E_1 for the electron in the potential box, has become split into two by the presence of the other box adjacent to it. These *two levels are equally spaced about the original level*, and, as is evident from equation (4), the difference between each level and the unaltered level is given by

$$\Delta E = \frac{hE_1}{\pi a\sqrt{2\mu(U - E_1)}} \cdot \epsilon^{-\frac{4\pi b\sqrt{2\mu(U-E_1)}}{h}}. \tag{6}$$

It will be observed that the exponential factor represents the *probability of penetration of the electron* through the barrier between the two potential boxes. This is evident from the remarks in Chapter III, especially those relating to equation (3.44). Thus for b, very large, or $U - E$, infinitely great, the value of this probability becomes infinitesimally small, and ΔE tends to vanish.

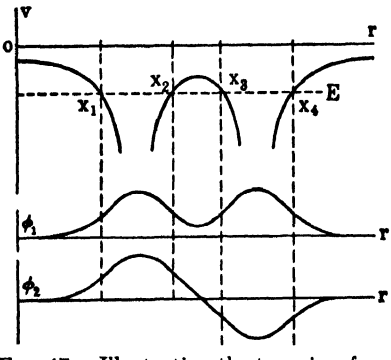


FIG. 47. Illustrating the two eigenfunctions for electron in ionized hydrogen molecule (H_2^+).

The bearing of these conclusions on the problem of the energy levels for the *ionized hydrogen molecule* (H_2^+) is self-evident. The two potential barriers are represented by the potential fields around the two nuclei, and the single electron must fluctuate between these two coulomb fields. The potential function $(-e^2/r)$ is indicated in Fig. 47[10], by the two sets of hyperbolas which are joined in the center. The energy of the electron in the lowest state is given for the individual atom, as shown

[10] N. F. Mott, *loc. cit.*, Fig. 10, p. 96.

in Chapter VII, by the relation,

$$E = -\frac{2\pi^2\mu e^4}{h^2} = -\frac{e^2}{2a_0},$$

and the eigenfunction is given by

$$\phi = \frac{1}{\sqrt{\pi}}\left(\frac{1}{a_0}\right)^3 \cdot \epsilon^{-\frac{2r}{a_0}}.$$

But when the two nuclei are brought together within a distance comparable with a_0, the probability of transition of the electron from one nucleus to the other becomes appreciable, and consequently each energy level becomes split up into two levels, very close together and only slightly different, of which E_S is less than E, that is *more negative*, and E_A lies above E, i.e., is more positive. Also corresponding to these two energy levels into which each original level is split, there are two separate modes of vibration, or eigenfunctions, which are indicated in Fig. 47 by ϕ_1 (corresponding to E_S) and ϕ_2 (corresponding to E_A).

8.4 Coupled Linear Harmonic Oscillators. The behavior of an electron in two neighboring potential boxes is represented, as discussed in the previous sections, by two modes of vibration. Two identical linear harmonic oscillators when coupled together exhibit a similar behavior, and it was by investigating the nature of the interaction of two such oscillators that F. London deduced his theory of the origin of van der Waals forces. Before presenting this theory, however, it is necessary to review briefly the concepts, electric moment, and polarizability of molecules.[11]

An atom of argon or molecule of methane has a symmetrical distribution of electrons about the nucleus or nuclei. That is, the center of gravity of the electrons coincides with that of the positive charges. Such molecules are said to be non-polar. If, however, these molecules are placed in an electric field as, for instance, that produced between the plates of a charged condenser, the positive and negative charges in the molecule are attracted toward the oppositely charged plates, so that the two centers of gravity no longer coincide. If z denote the separation of the two sets of charges, the molecule now behaves like a dipole of electric moment $\mu_e = ez$. The magnitude of z, and consequently that of the induced moment, is proportional to the field F (in volts per centimeter), in accordance with the relation

$$\mu_e = ez = \alpha F, \tag{7}$$

[11] See references at the end of the chapter.

and the constant α is designated the *polarizability.* The *polarization energy* is therefore given by

$$U_p = -\int_0^{\mu_e} F d\mu_e = -\int_0^F \alpha F dF = -\frac{\alpha F^2}{2}, \tag{8}$$

where the negative sign indicates that the energy is that of attraction.

In order to illustrate London's theory in the simplest manner, let us consider the interaction of two linear harmonic oscillators, arranged as shown in Fig. 48, where R is the equilibrium distance between the positive ends of the dipoles, and u_1 and u_2, representing the amplitudes of oscillation, are each small compared to R.[12] We shall assume that the masses μ and charge e are the same for both oscillators, as well as the restoring force constant k.

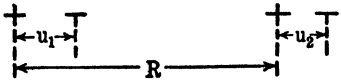


FIG. 48. Interaction of two dipoles arranged in the same direction.

The mutual potential energy of the two oscillators is given, according to Coulomb's law, by the expression

$$U = e^2\left\{\frac{1}{R+u_2-u_1} - \frac{1}{R+u_2} - \frac{1}{R-u_1} + \frac{1}{R}\right\}$$

$$= \frac{e^2}{R}\left\{\frac{1}{1+\dfrac{u_2-u_1}{R}} - \frac{1}{1+\dfrac{u_2}{R}} - \frac{1}{1-\dfrac{u_1}{R}} + 1\right\}.$$

Since u_1/R and u_2/R are each considerably smaller than 1, we can expand each of the terms thus:

$$\left(1+\frac{u_2}{R}\right)^{-1} = 1 - \frac{u_2}{R} + \left(\frac{u_2}{R}\right)^2 - \left(\frac{u_2}{R}\right)^3 + \ldots$$

$$\left(1-\frac{u_1}{R}\right)^{-1} = 1 + \frac{u_1}{R} + \left(\frac{u_1}{R}\right)^2 + \left(\frac{u_1}{R}\right)^3 + \ldots$$

and

$$\left\{1+\frac{u_2-u_1}{R}\right\}^{-1} = 1 - \frac{u_2-u_1}{R} + \frac{u_2^2+u_1^2-2u_2u_1}{R^2} - \ldots$$

[12] This illustration is taken from the discussion by M. Born and M. Goppert-Mayer in " Handbuch der Physik," Vol. XXIV, Part 2, p. 750.

Neglecting terms in these expansions involving R^{-3} and more negative powers of R, we obtain the relation

$$U = -\frac{2e^2u_1u_2}{R^3}. \tag{9}$$

By breaking the series at this point we take into account only dipole-dipole interaction. The following terms of the series would correspond to dipole-quadrupole interaction (term involving R^{-4}) and quadrupole-quadrupole interaction (term involving R^{-5}).

We can now write down the total energy of the system in the Hamiltonian form (see Chapter IV)

$$E = H(p, q) = \frac{1}{2\mu}(p_1^2 + p_2^2) + \frac{k}{2}(u_1^2 + u_2^2) - \frac{2e^2u_1u_2}{R^3}. \tag{10}$$

For large values of R the last term vanishes, and we have two identical linear oscillators, for which the frequency is given by the relation

$$\nu_0 = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}. \tag{11}$$

When, however, R is decreased, the last term in equation (10) has to be taken into account. The field due to each oscillator acts on the other, and, in accordance with equation (8),

$$\tfrac{1}{2}ku^2 = \tfrac{1}{2}\alpha F^2, \tag{12}$$

where F is the field and u the resultant displacement.

The resulting dipole moment is given by

$$\mu_e = ue = \alpha F. \tag{13}$$

It follows from (12) and (13) that

$$\alpha = \frac{e^2}{k}. \tag{14}$$

In the absence of the last term in equation (10) the total energy E is the sum of the energies E_1 and E_2 of two simple oscillators. However, the presence of the product term involving u_1u_2, which is designated the *perturbation potential energy term*, shows that the simple additive solution cannot be valid. Under these circumstances it is necessary in general to resort to the perturbation theory, and in wave mechanics there is available a similar mathematical technique for dealing with the corresponding S. equation, which is discussed in the following chapter. In equation (10) it is possible, however, to obtain an expression for E by a much less tedious procedure. This is the method of transformation to so-called *normal coördinates* (" Hauptachsen ").

In equation (10) let us introduce the new coördinates q_1 and q_2 such that

$$\left.\begin{aligned} q_1 &= \frac{1}{\sqrt{2}}(u_1 + u_2), \\ \text{and} \quad q_2 &= \frac{1}{\sqrt{2}}(u_1 - u_2). \end{aligned}\right\} \tag{15}$$

$$\left.\begin{aligned} \text{Hence,} \quad u_1 &= \frac{1}{\sqrt{2}}(q_1 + q_2), \\ \text{and} \quad u_2 &= \frac{1}{\sqrt{2}}(q_1 - q_2). \end{aligned}\right\} \tag{16}$$

It will be recognized that this *transformation corresponds to a rotation* of the rectangular axes u_1 and u_2 through an angle of 45°. This accounts for the similarity between the two sets of equations (15) and (16).

In terms of these new variables

$$\begin{aligned} p_1^2 + p_2^2 &= \mu^2 \{\dot{u}_1^2 + \dot{u}_2^2\} \\ &= \mu^2 \{\dot{q}_1^2 + \dot{q}_2^2\}, \end{aligned}$$

so that p_1^2 and p_2^2 now refer to momenta associated with the new coördinates.

Furthermore,

$$2u_1u_2 = q_1^2 - q_2^2.$$

Therefore equation (10) becomes

$$E = \frac{1}{2\mu}(p_1^2 + p_2^2) + \left(\frac{k}{2} - \frac{e^2}{R^3}\right)q_1^2 + \left(\frac{k}{2} + \frac{e^2}{R^3}\right)q_2^2. \tag{17}$$

In this form, the energy is again separable into two terms, corresponding to *two new modes of vibration* of which the frequencies ν_1 and ν_2 are given by the relations

$$\left.\begin{aligned} \nu_1 &= \frac{1}{2\pi}\sqrt{\frac{k}{\mu}\left(1 - \frac{2e^2}{kR^3}\right)}; \\ \nu_2 &= \frac{1}{2\pi}\sqrt{\frac{k}{\mu}\left(1 + \frac{2e^2}{kR^3}\right)}. \end{aligned}\right\} \tag{18}$$

It is also evident from the equations in (15) that ν_1 and ν_2 correspond to the symmetric and antisymmetric modes respectively.

For values of $x < 1$

$$\left.\begin{aligned}(1-x)^{\frac{1}{2}} &= 1 - \frac{x}{2} - \frac{x^2}{8} - \frac{x^3}{16} - \ldots \\ \text{and} \qquad (1+x)^{\frac{1}{2}} &= 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \ldots .\end{aligned}\right\} \tag{19}$$

Since $2e^2/(kR^3)$ is small compared to unity, we can expand the expressions under the radical signs in (18). The resulting relations are given by

$$\nu_1 = \nu_0\left(1 - \frac{2e^2}{kR^3} - \frac{e^4}{2k^2R^6} + \ldots\right),$$

and

$$\nu_2 = \nu_0\left(1 + \frac{2e^2}{kR^3} - \frac{e^4}{2k^2R^6} + \ldots\right),$$

where ν_0 is given by equation (11).

For a single oscillator, in accordance with the result obtained in Chapter V,

$$E_n = h\nu_0(n + \tfrac{1}{2}). \tag{5.16}$$

Hence, in the present case

$$\begin{aligned}E_{n_1n_2} &= h\nu_1\,(n_1 + \tfrac{1}{2}) + h\nu_2(n_2 + \tfrac{1}{2}) \\ &= h\nu_0\left\{n_1 + n_2 + 1 - \frac{2e^2}{kR^3}(n_1 - n_2)\right. \\ &\quad \left. - \frac{e^4}{2k^2R^6}(n_1 + n_2 + 1) + \ldots\right\}.\end{aligned}$$

For $n_1 = n_2 = 0$,

$$\begin{aligned}E_{00} &= h\nu_0\left(1 - \frac{e^4}{2k^2R^6}\right) \\ &= h\nu_0\left(1 - \frac{\alpha^2}{2R^6}\right)\end{aligned}$$

in consequence of (14).

Thus the *coupling energy* ΔE is given by the relation

$$\Delta E = -\frac{h\nu_0\alpha^2}{2R^6}. \tag{20}$$

Before treating the same problem from the point of view of the S. equation, it is of interest to consider the relatively simpler case of a linear harmonic oscillator acted on by a uniform electric field F.[13]

In absence of a field, as shown in Chapter V, the S. equation has the form

$$\frac{d^2\phi}{dq^2} + a\left(E - \frac{kq^2}{2}\right)\phi = 0, \tag{21}$$

where $a = 8\pi^2\mu/h^2$.

The eigenfunction corresponding to the state $n = 0$ is

$$\phi_0 = \left(\frac{1}{\pi}\right)^{\frac{1}{4}} \epsilon^{-\frac{x^2}{2}},$$

where $x = 2\pi q\sqrt{\frac{\mu\nu_0}{h}} = q\sqrt{b}$.

Since

$$1 = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} \epsilon^{-x^2} dx = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} \epsilon^{-bq^2}\sqrt{b} \cdot dq,$$

therefore, ϕ_0, as a function of the actual displacement q, is given by

$$\phi_0(q) = \left(\frac{b}{\pi}\right)^{\frac{1}{4}} \epsilon^{-\frac{bq^2}{2}}.$$

Let

$$c = \frac{2\pi^2\mu}{h}.$$

Then,

$$b = \frac{4\pi^2\mu\nu_0}{h} = 2c\nu_0,$$

and

$$\phi_0(q) = \left(\frac{2c\nu_0}{\pi}\right)^{\frac{1}{4}} \epsilon^{-c\nu_0 q^2}. \tag{22}$$

If the oscillator is subject to a *uniform electric field* F the potential energy term in (21) becomes

$$V = (\tfrac{1}{2})kq^2 - eFq$$
$$= \frac{k}{2}\left(q - \frac{eF}{k}\right)^2 - \frac{e^2F^2}{2k}.$$

[13] The following discussion is based upon the paper by J. E. Lennard-Jones, *Proc. Phys. Soc.* (*London*), **43**, 461 (1931).

Thus, if we let $z = q - eF/k$, the wave equation (21) becomes

$$\frac{d^2\phi}{dz^2} + a\left(E + \frac{e^2F^2}{k^2} - \frac{kz^2}{2}\right)\phi = 0, \tag{23}$$

which is of the same form as (21). The energy values, however, are given by the relation

$$E = (n + \tfrac{1}{2})h\nu_0 - \left(\frac{e^2F^2}{2k}\right). \tag{24}$$

That is, the value of ν_0 is unaltered, and $\phi_0(z)$ is identical in form with $\phi_0(q)$, but the whole wave pattern is *displaced* a distance eF/k in the direction of the field. This is similar to the phenomenon which occurs when a molecule or atom is polarized, and the *polarization energy* is given, according to the last equation, by

$$E_p = -\frac{e^2F^2}{2k} = -\frac{\alpha F^2}{2}, \tag{25}$$

where $\alpha = e^2/k$.

Let us now consider *two identical linear oscillators*, each vibrating along the q-axis, with centers a distance R apart. There will be an interaction due to the field produced by each oscillator in the region of the other, and in accordance with equation (10), the corresponding S. equation has the form

$$\frac{\partial^2\phi}{\partial q_1^2} + \frac{\partial^2\phi}{\partial q_2^2} + a\left(E - \frac{kq_1^2}{2} - \frac{kq_2^2}{2} + \frac{2e^2q_1q_2}{R^3}\right)\phi = 0, \tag{26}$$

where ϕ is a function of q_1 and q_2 as well as of R.

This is a type of equation which occurs very often in quantum mechanics, in fact *it is typical of all cases in which there is an interaction between two systems.* We meet with it, in a modified form, in the investigation of the helium atom and of the hydrogen molecule.

Now if the perturbation energy term were not present, the solution of the equation

$$\frac{\partial^2\phi}{\partial q_1^2} + \frac{\partial^2\phi}{\partial q_2^2} + a\left(E - \frac{kq_1^2}{2} - \frac{kq_2^2}{2}\right)\phi = 0 \tag{27}$$

would obviously be obtained by putting

$$\phi(q_1, q_2) = \phi(q_1) \cdot \phi(q_2).$$

For, if we substitute the latter in (27) we obtain the equation

$$\phi(q_2)\left\{\frac{d^2\phi(q_1)}{dq_1^2}+\frac{a}{2}(E-kq_1^2)\phi(q_1)\right\} \quad + \quad \phi(q_1)\left\{\frac{d^2\phi(q_2)}{dq_2^2}+\frac{a}{2}(E-kq_2^2)\phi(q_2)\right\}=0.$$

Since $\phi(q_1)$ is a function of q_1 only, and $\phi(q_2)$ of q_2 only, it follows that each of the expressions in large brackets is equal to zero. That is, we obtain the two ordinary differential equations which are each similar to the S. equation for the linear harmonic oscillator. Consequently, the solutions are similar, with the only difference that for q_1

$$E_1 = (n_1 + \tfrac{1}{2})h\nu_0,$$

and for q_2

$$E_2 = (n_2 + \tfrac{1}{2})h\nu_0,$$

where $E = E_1 + E_2$ = total energy and n_1 and n_2 are not necessarily identical.

In order to solve equation (26) we use the same transformation to normal coördinates as indicated by equations (15) and (16). That is, we set

$$\left.\begin{aligned} z_1 &= \frac{1}{\sqrt{2}}(q_1+q_2) \\ z_2 &= \frac{1}{\sqrt{2}}(q_1-q_2) \end{aligned}\right\}, \tag{28}$$

and, as in the derivation of equation (17), we obtain the S. equation in the form

$$\frac{\partial^2\phi}{\partial z_1^2}+\frac{\partial^2\phi}{\partial z_2^2}+a\left(E-\frac{k_1z_1^2}{2}-\frac{k_2z_2^2}{2}\right)\phi=0, \tag{29}$$

where

$$\left.\begin{aligned} k_1 &= k-\frac{2e^2}{R^3} \\ \text{and} \quad k_2 &= k+\frac{2e^2}{R^3}. \end{aligned}\right\} \tag{30}$$

Equation (29) is evidently separable into two ordinary differential equations, of which the solutions are $\phi_{n_1}(z_1)$ and $\phi_{n_2}(z_2)$ with the eigenvalues E_{n_1} and E_{n_2}, and the two new frequencies are given by equations (18).

It follows from equations (30) that the frequency ν_1, which is less than ν_0, corresponds to that of the *symmetric mode*, and ν_2 (which is greater than ν_0) corresponds to the frequency of vibration for the *antisymmetric mode*. Thus the behavior of two coupled oscillators resembles that of the electron in two potential boxes or in the ionized hydrogen molecule. Instead of an energy $E_1 = h\nu_0$ for the lowest state of the oscillators, we obtain the result $E = (\frac{1}{2})h\nu_1 + (\frac{1}{2})h\nu_2$.

Obviously, the transformation to principal coördinates is no longer valid if k is less than $2e^2/R^3$. That is, the conclusion involved in (20) is tenable only if $R^3 > (2e^2/k)$, which determines a *lower limit* for the distance R between the zero positions of the oscillators. Since, as shown in equation (14), $e^2/k = \alpha$, the *polarizability*, we can also state this limit in the form $R^3 > 2\alpha$. It should be observed that α may be derived from measurements of dielectric constant or refractive index, and is of the order of magnitude 10^{-24} cm.[3] Thus this theory of interaction of linear oscillators is valid for values of $R > 10^{-8}$ cm. approximately, which corresponds to intermolecular distances.

For the state of lowest energy $n_1 = n_2 = 0$, we thus obtain the value for the coupling energy ΔE given by equation (20).

The negative sign shows that the energy arises from *attractive forces*, and, since the energy varies as R^{-6}, the force of attraction is given by

$$F_A = -\frac{dE}{dR} = -\frac{3\alpha^2 h\nu_0}{R^7}. \tag{31}$$

That is, the *force of attraction between two pulsating dipoles varies inversely as the seventh power of the distance.* It, therefore, falls off rapidly with increase in distance. Thus, if R is doubled, the force decreases to $(\frac{1}{2})^7 = 1/128$ of its original value.

The behavior of the eigenfunction $\phi(z_1, z_2)$ throws additional light upon the nature of this attractive force between linear oscillators. For the lowest state ($n_1 = n_2 = 0$), it follows from equations (22) and (29) that

$$\phi(z_1, z_2) = \phi(z_1) \cdot \phi(z_2)$$

$$= \left(\frac{4c^2\nu_1\nu_2}{\pi^2}\right)^{\frac{1}{4}} \epsilon^{-c(\nu_1 z_1^2 + \nu_2 z_2^2)}, \tag{32}$$

$$= \left(\frac{\nu_1\nu_2}{\nu_0^2}\right)^{\frac{1}{4}} \{\phi_0(q)\}^2 \epsilon^{-\left(\frac{c\nu_0}{k}\right)V}, \tag{33}$$

where $\{\phi_0(q)\}^2$ is the product $\phi(q_1) \cdot \phi(q_2)$ for $R = \infty$ [solution of equation (27)] and V is the perturbing potential energy function,

defined by the relation

$$V = -\frac{2e^2 z_1 z_2}{R^3}.$$

From equation (33) we deduce the probability distribution function for the coördinates of the two oscillators in the form

$$\phi^2(z_1, z_2) = \left(\frac{\nu_1 \nu_2}{\nu_0^2}\right)^{\frac{1}{2}} \{\phi_0(q)\}^4 \epsilon^{-\frac{2c\nu_0}{k}V}. \tag{34}$$

The exponent of ϵ in this equation is given by the expression $4c\nu_0 \alpha z_1 z_2/R^3$, which shows that when z_1 and z_2 are of the *same sign* (the pulsations are *in phase*), $\phi^2(z_1, z_2) > \{\phi(q_1) \cdot \phi(q_2)\}^2$, and when z_1 and z_2 are opposite in phase, the reverse is true. As Lennard-Jones points out:

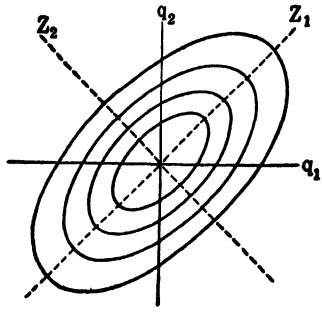


FIG. 49. Phase relations for two coupled linear harmonic oscillators.

> This result can be represented diagrammatically in a two-dimensional space, as shown in Fig. 49. The distribution function, being proportional to $\epsilon^{-2c(\nu_1 z_1^2 + \nu_2 z_2^2)}$ [see equation (32)], is like a ridge with its maximum at the origin, and with elliptical contours.[14] The major axes of the ellipses are $z_2 = 0$, or $q_1 = +q_2$, and the minor axes $q_1 = -q_2$. The corresponding distribution without interaction is a round-topped mountain with circular contours. The probability of finding q_1 and q_2 with the same sign has increased, while that of finding them with opposite sign has decreased. In other words, the interacting dipoles tend on the average to move in phase.

8.5 Van der Waals Interaction Energy for Two Hydrogen Atoms. Let us now consider the interaction of two hydrogen atoms separated by a distance R which is *large* compared with the average distance of the electron in each atom from its associated nucleus. (The following argument is valid only for this case.) Disregarding the potential energy terms due to the force between electron and nucleus in each atom, the

[14] The equation of an ellipse with respect to rectangular coördinates through the center is $x^2/a^2 + y^2/b^2 = 1$, where a and b are semi-major and -minor axes, respectively. Hence the expression $2c(\nu_1 z_1^2 + \nu_2 z_2^2)$ is constant for all values of z_1 and z_2 which occur on the ellipse whose semi-axes are $1/\sqrt{2c\nu_1}$ and $1/\sqrt{2c\nu_2}$. Thus, if we have a series of confocal ellipses, as indicated in Fig. 49, we can describe any simultaneous combination of values of z_1 and z_2 by stating the serial number of the corresponding ellipse. The use of such coördinates is often very convenient in mathematical problems and has proved useful in dealing with problems of potential.

potential energy arising from the Coulomb forces between charged particles in one atom (nucleus A and electron 1) and similar particles in the other (nucleus B and electron 2) is given (see Fig. 50) by

$$V = e^2\left\{\frac{1}{R} + \frac{1}{r_{12}} - \frac{1}{r_{B1}} - \frac{1}{r_{A2}}\right\}. \qquad (35)$$

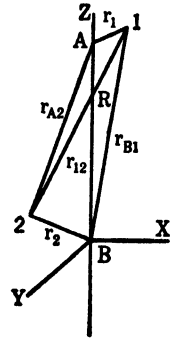


FIG. 50. Illustrating van der Waals interaction of two hydrogen atoms.

Let $x_1y_1z_1$ and $x_2y_2z_2$ designate the coördinates of each electron (1 and 2 respectively) with respect to its nucleus (A and B respectively), and let the line joining the nuclei lie along the z-axis. Then

$$\begin{aligned} r_{12}^2 &= (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 + R - z_2)^2, \\ r_{A2}^2 &= x_2^2 + y_2^2 + (R - z_2)^2, \\ r_{B1}^2 &= x_1^2 + y_1^2 + (R + z_1)^2, \\ r_1^2 &= x_1^2 + y_1^2 + z_1^2, \\ r_2^2 &= x_2^2 + y_2^2 + z_2^2. \end{aligned}$$

Hence,

$$\begin{aligned} r_{12}^2 &= r_1^2 + r_2^2 - 2(x_1x_2 + y_1y_2 + z_1z_2) + 2R(z_1 - z_2) + R^2, \\ r_{A2}^2 &= R^2 + r_2^2 - 2Rz_2, \\ r_{B1}^2 &= R^2 + r_1^2 + 2Rz_1. \end{aligned}$$

Also

$$\frac{1}{r_{12}} = \frac{1}{R}\left\{1 + \left(\frac{r_1}{R}\right)^2 + \left(\frac{r_2}{R}\right)^2 - \frac{2}{R^2}(x_1x_2 + y_1y_2 + z_1z_2) + \frac{2}{R}(z_1 - z_2)\right\}^{-\frac{1}{2}},$$

and similar expressions may be written for $1/r_{A2}$ and $1/r_{B1}$. Replacing each of the expressions in the brackets as in (18), by means of the expansion

$$(1 - x)^{-\frac{1}{2}} = 1 + \frac{x}{2} + \frac{3x^2}{8} + \frac{5x^3}{16} + \ldots,$$

we obtain the relation

$$V = \frac{e^2}{R^3}\{x_1x_2 + y_1y_2 + z_1z_2 + \tfrac{3}{2}(z_1 - z_2)^2 - \tfrac{3}{2}(z_1^2 + z_2^2)\} + \text{terms involving } (1/R)^4 \text{ and higher powers of } (1/R).$$

The expression involving $1/R^3$ corresponds to the dipole-dipole interaction and is given by

$$V_I = \frac{e^2}{R^3}(x_1x_2 + y_1y_2 - 2z_1z_2). \tag{36}$$

Equation (36) may be derived directly,[15] without going through the preceding argument, if we consider the motion of the electron in each atom as equivalent to that of a dipole, for which both the electric moment and direction of orientation are varying continuously. From this point of view we therefore regard each atom as a dipole for which the respective components along the three coördinate axes are ex_1, ey_1, ez_1, and ex_2, ey_2, ez_2. The result derived in equation (36) is thus an extension of that derived in equation (9).

Lennard-Jones writes:

> The probability distribution function can now be represented only in six-dimensional space. It appears that z_1 and z_2 tend to have the same sign (as before), while x_1 tends to be opposite to x_2, and y_1 opposite to y_2. These are just the configurations for which the oscillators attract. Hence, we may say that two systems tend to interact in such a way that the attraction is a maximum. It is in the generalized sense that we use the expression, " tend to move in phase," for actually in this example the x's and y's tend to be out of phase in the usual sense.

Such a procedure fails, however, to take into account terms involving powers of $(1/R)$ higher than $(1/R)^3$. Furthermore, as will be shown in the following chapter, it is possible, by application of the second-order perturbation theory to the relation for V given in equation (35), to derive a more accurate value of the interaction energy of two atoms.

It follows that the attraction due to each term x_1x_2, y_1y_2, and z_1z_2 can be treated separately, and as shown by H. R. Hassé,[16] the interaction energy due to each of the first two terms is $\frac{1}{4}$ of that due to the z_1 and z_2 oscillators, and therefore equal to $-\alpha^2h\nu_0/(8R^6)$. Hence, the total interaction energy for a pair of three-dimensional oscillators is

$$\Delta E = -\frac{3}{4}\frac{\alpha^2h\nu_0}{R^6}. \tag{37a}$$

Let us now attempt to apply this result to the calculation of the *interaction energy for two hydrogen atoms.*

[15] See, for instance, J. H. Jeans' " The Mathematical Theory of Electricity and Magnetism," Cambridge Univ. Press (1908), p. 368, equation (355).

[16] Hassé, *Proc. Cambridge Phil. Soc.*, **27**. 66 (1931). See also F. London, *Z. physik. Chemie*, **B11**, 222 (1931).

By investigating the behavior of a Bohr orbit in an electric field it has been shown that the polarizability of the atom is given by $\alpha = (9/2)a_0^3$, where a_0 is the radius of the orbit.[17]

Hence the interaction energy of two hydrogen atoms is given by the relation

$$\Delta E = -\frac{243}{16} \cdot \frac{h\nu_0}{R^6}. \tag{37b}$$

What value of ν_0 can we assign to the electron in a hydrogen atom? Obviously, this "frequency" is a fiction, since, on the basis of the Principle of Indeterminacy, no experiment can be devised to determine a frequency of revolution of the electron in its orbit. But when in doubt, it is customary to appeal to the Principle of Correspondence (see Chapter I). According to this principle the behavior of the electrons in an atomic system must approach more and more that predicted by classical physics the higher the quantum number of the orbit. This is equivalent to the statement that when dealing with larger-scale phenomena the results deduced by quantum mechanical methods must approximate those deduced by classical mechanics.

In the case of the hydrogen atom, the electronic orbits increase in radius as the ionizing stage is approached. If V_i denote the ionization potential, then the limit of the line spectrum, that is, the maximum frequency of light which can be emitted by the return of an electron to the lowest level, is given by

$$\nu_0 = \frac{V_i e}{h}. \tag{38}$$

According to electromagnetic theory, radiation of this frequency would be emitted by an oscillator having the identical frequency ν_0. Therefore, we might use in equation (37a) the value of ν_0 given by the last equation.

Again, the resonance potential (V_r) of hydrogen corresponds to the first excited state, and the frequency of radiation emitted because of a transition from this level to the lowest or normal state is

$$\nu_0 = \frac{V_r e}{h}. \tag{39}$$

Since $V_r = V_i(1 - \frac{1}{4}) = (\frac{3}{4})V_i$, the value of ν_0 is thus confined to the two limits defined by equations (38) and (39).

[17] See P. Debye, "Polar Molecules," Chapter I.

Now the energy of the hydrogen atom in the normal state is given by[18]

$$E_0 = -\frac{e^2}{2a_0}.$$

Since $V_{ie} = -E_0$, the corresponding frequency is

$$\nu_0 = \frac{e^2}{2a_0 h}. \tag{40}$$

Substituting for ν_0 in (37*b*), the interaction energy is found to be given by a relation of the form

$$\Delta E = -\frac{Ce^2a_0^5}{R^6}, \tag{41}$$

where $C = 243/32 = 7.59$ if ν_0 is calculated by means of equation (38), and $C = \frac{3}{4} \times 7.59 = 5.69$ if equation (39) is used.

The result thus obtained obviously represents only a first approximation to the correct value. A calculation of the second-order perturbation energy, in which the perturbing potential term is given by (36), yields a value[19] $C = 6$. For a more accurate calculation it is necessary to take into consideration dipole-quadrupole and quadrupole-quadrupole interaction, as has been done for helium by H. Margenau.[20] The calculation can also be made by using the method of variation of parameters, which is discussed in a subsequent chapter.

Such calculations have been carried out by R. Eisenschitz and F. London,[21] H. R. Hassé,[22] J. E. Lennard-Jones,[23] J. C. Slater and J. G. Kirkwood,[24] and L. Pauling and J. Y. Beach.[25] The values derived for C by the different investigators are shown in the following summary:

Hassé	6.49
Slater and Kirkwood	6.49
Pauling and Beach	6.4984; 6.499

The value $C = 6.499$, obtained by Pauling and Beach, is probably "extremely close to the correct value."[26] Inserting this value of C

[18] $E_0 = -13.53$ v.e.
[19] Pauling and Wilson, "Introduction to Quantum Mechanics," p. 385.
[20] *Phys. Rev.*, **38**, 747 (1931).
[21] *Z. Physik*, **60**, 491 (1930).
[22] *Proc. Cambridge Phil. Soc.*, **27**, 66 (1931).
[23] *Proc. Roy. Soc.* (*London*), **A129**, 598 (1930).
[24] *Phys. Rev.*, **37**, 682 (1931).
[25] *Phys. Rev.*, **47**, 686 (1935).
[26] Pauling and Wilson, "Introduction to Quantum Mechanics," p. 382.

and the values $e = 4.770 \times 10^{-10}$ e.s.u., $a_0 = 0.528 \times 10^{-8}$ cm., the value of the interaction energy of two hydrogen atoms becomes

$$\Delta E = -6.064 \times 10^{-60} R^{-6} \text{ erg.} \tag{42}$$

8.6 Van der Waals Energy for Other Atoms and Derivation of Constant "a." For the case of two helium atoms, F. London's[27] method of calculating the energy of interaction leads to the relation

$$(\Delta E)_2 = -\frac{3}{4}\frac{\alpha^2 V_i}{R^6},$$

where α is the polarizability and V_i the ionization potential. Substituting the value $\alpha = 0.205 \times 10^{-24}$ which is that calculated from the refractive index, and the value $V_i = 24.5$ volts, this corresponds to

$$(\Delta E)_2 = -2.65 E_0 \left(\frac{a_0}{R}\right)^6,$$

where $E_0 = \dfrac{e^2}{2a_0}$.

By application of the perturbation theory, Slater and Kirkwood[28] deduced for this interaction energy the value

$$(\Delta E)_2 = -3.18 E_0 \left(\frac{a_0}{R}\right)^6$$

$$= -1.59 e^2 a_0^5 \cdot R^{-6}.$$

For more complex atoms and molecules Slater and Kirkwood also deduced for the energy of interaction a relation of the form

$$(\Delta E)_3 = -1.36 v^{\frac{1}{2}} a_0^{\frac{3}{2}} \alpha^{\frac{3}{2}} E_0 \cdot R^{-6}, \tag{44a}$$

where v = number of electrons in valence shell.

While this equation, as well as (42) and (43), expresses the energy in terms of *ergs per molecule*, it is of interest to calculate the magnitude of this energy per gram-molecule. In terms of kg.cal./mole,

$$(\Delta E)_3 = -163 \times \frac{10^{-12} v^{\frac{1}{2}} \alpha^{\frac{3}{2}}}{R^6}. \tag{44b}$$

Thus, in the case of CH_4, $v = 8$, $\alpha = 2.59 \times 10^{-24}$, and van der Waals' constant $b = 2\pi N R^3/3 = 55.9$ cm.^{3}/mole. Hence $(\Delta E)_3 = 991.2$ cal./mole, that is, approximately 1000 cal./mole. Since this value is of the same order of magnitude as the heat of evaporation of

[27] *Z. physik. Chemie*, **B11**, 222 (1931).

[28] *Phys. Rev.*, **37**, 682 (1931).

solid or liquid methane, the result obtained indicates that the attractive forces which give rise to cohesion in non-polar liquids and solids are probably of the same nature as those due to the interaction of fluctuating dipoles. As will be shown in a subsequent section, a more rigorous calculation confirms this conclusion.

Referring now to equation (iv) in section 1, it is evident that

$$\Delta E = -\frac{A}{(m-1)r^{m-1}}.$$

Since $m - 1 = 6$, as deduced on the basis of London's theory, we obtain the relation

$$A = -6 \cdot \Delta E \cdot R^6.$$

Thus, for the case of hydrogen atoms, it follows from equation (42) that $A_{\mathrm{H}} = 36.386 \times 10^{-60}$ dyne $\cdot$ cm.^{6}, while for the case of helium atoms, it follows from equation (43) that

$$A_{\mathrm{He}} = 8.91 \times 10^{-60} \text{ dyne} \cdot \text{cm.}^6.$$

London's theory would lead to the lower value

$$A_{\mathrm{He}} = 7.44 \times 10^{-60} \text{ dyne} \cdot \text{cm.}^6.$$

For other atoms and molecules values of A may be derived by means of equation (44a).

In order to calculate a it is necessary, as is evident from equation (vii), to determine the repulsive energy term involving B in equation (iv). J. E. Lennard-Jones has shown that the magnitude of this term may be derived empirically from PV versus T data for the gases, on the basis $m = 7$. The values of n thus derived vary from 10 to 13, indicating that the repulsive force increases extremely rapidly with decrease in r below the equilibrium value r_0.

J. C. Slater and J. G. Kirkwood[29] have shown that, in the case of helium, the repulsive potential energy term may be written in the form

$$U(\text{repulsive}) = B\epsilon^{-\frac{cr}{a_0}},$$

where B and c are constants. A similar expression has been deduced by M. Born and J. E. Mayer[30] for the repulsive energy between two ions in a halide lattice. Actually, it makes little difference whether the exponential form is used or an expression of the form r^{-n}, if n is as large

[29] *Phys. Rev.*, **37**, 682 (1931).
[30] *Z. Physik*, **75**, 1 (1932).

as 10. Figure 51[31] gives the potential energy of pairs of inert gas atoms as a function of their distance apart in Ångström (10^{-8} cm.), and it will be observed that, owing to the extremely rapid increase in the repulsive force with decrease in distance beyond the equilibrium value (the value of r at the minimum), $U(r)$ increases rapidly in this range. The force at any value of r is, of course, determined by the slope at that point.

For a first approximation we can assume that the value of n, the repulsive force exponent, is so large that the repulsive energy term in equation (iv) may be neglected. Hence, we can write equation (vii) in the form

$$a = -\frac{2\pi N^2}{1.013 \times 10^6}\int_{r_0}^{\infty} \Delta E \cdot r^2 dr, \quad (45)$$

where ΔE is the attractive energy term calculated by means of equations (42), (43), or (44). Expressing the latter in the form

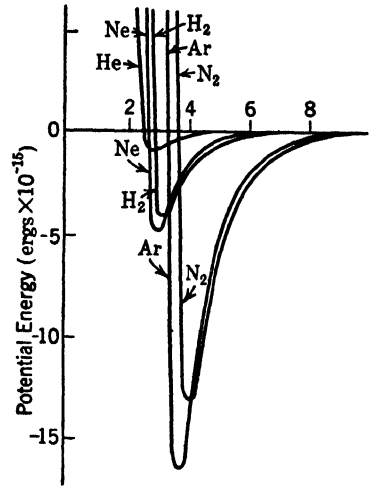


FIG. 51. Van der Waals potential energy curves for several gases.

$$\Delta E = -A_0 \cdot r^{-6}$$

we obtain the relation

$$\int_{r_0}^{\infty} \Delta E \cdot r^2 dr = -A_0 \int_{r_0}^{\infty} \frac{dr}{r^4} = -\frac{A_0}{3r_0^3}.$$

Since van der Waals' constant b is given in terms of r_0^3 by the relation

$$r_0^3 = \frac{3b}{2\pi N},$$

we can write (45) in the form

$$a = \frac{\pi^2 N^3 A_0}{3.039 \times 10^6 b} = \frac{9.654 \times 10^{65} A_0}{b}. \quad (46)$$

Hence, in terms of equation (44)

$$a = \frac{1.08 \times 10^{43} \alpha^{\frac{3}{2}} v^{\frac{1}{2}}}{b}, \quad (47)$$

[31] J. E. Lennard-Jones, *Proc. Phys. Soc.* (*London*), **43**, 461 (1931).

and in terms of London's relation

$$a = \frac{1.13 \times 10^{54} \alpha^2 V}{b}, \tag{48}$$

where V is the ionization or resonance potential (in volts).

Both these relations have been shown to give values of a which are in reasonably satisfactory agreement with the values derived from the critical constants by means of (viii).

8.7 Energies of Evaporation and Sublimation. The fact that the quantum mechanics theory leads to satisfactory agreement between calculated and observed values of van der Waals' constant a indicates that the attractive forces which give rise to cohesion in non-polar liquids and solids must be of the same nature as those described in the previous section.

As mentioned previously, the interaction energy per mole as calculated on the basis of London's theory is of the same order of magnitude as the energy of evaporation for solid CH_4. London and Lennard-Jones have shown that it is possible by a more exact application of the results deduced in the previous sections to calculate the energies of sublimation of crystal lattices of such molecules as those of the rare gases, methane, hydrogen, oxygen, nitrogen, and other gases. The present chapter would be incomplete without some discussion of this application of the quantum theory of van der Waals forces.

Given $U(r)$ for a pair of molecules it is possible to calculate the total potential energy of all the molecules in a gram-molecular volume of the crystal lattice. The method used may be described as a summation of the energy for every pair of molecules in the lattice. If we denote the total energy by ϕ_0, then

$$\phi_0 = \frac{N}{2}(C_A U_A + C_R U_R),$$

where C_A is the so-called *crystal potential constant*[32] for attraction, which depends upon both the value of the attractive force exponent m and the type of crystal lattice, and U_A is the attractive energy term. U_R designates the repulsive energy term and C_R is the corresponding crystal potential constant (which depends upon the value of n). The factor $(\frac{1}{2})$ is necessary to avoid counting the same atom twice.

The application of the last equation to calculate heats of evaporation may be illustrated by the results obtained for argon. F. London, in

[32] Values of these constants for different values of m and n and for various types of lattice structure, are given in Fowler's "Statistical Mechanics," Chapter X, also by J. E. Jones and A. E. Ingham, *Proc. Roy. Soc.* (*London*), **A107**, 636 (1925).

his calculation, assumes that the repulsive force exponent $n = \infty$, so that $U_R = 0$. For the attractive energy term U_A he uses the relation

$$U_A = -\frac{3}{4}\frac{\alpha^2 h\nu_0}{r_0^6} = 4.92 \times \frac{10^{-59}}{r_0^6}\text{erg},$$

where $h\nu_0 = V_i e$, with $V_i = 15.5$ electron volts, and $\alpha = 1.63 \times 10^{-24}$, while r_0 is the minimum distance between atoms.

The observed crystal spacing (minimum distance between atoms) is 3.84×10^{-8} cm., while $C_A = 115.4$. Hence, the energy of sublimation *in calories* is

$$S = \frac{\phi_0}{4.184 \times 10^7} = \frac{N}{2} \times \frac{4.92 \times 10^{-59} \times 115.4}{4.184 \times 10^7 \times (3.84 \times 10^{-8})^6}$$

$= 1715$ cal./mole.

The observed value, according to London, is 2030 cal./mole.

J. E. Lennard-Jones has carried out similar calculations in which the repulsive energy term has been taken into account. It should be mentioned, however, that, in general, the value of the term $C_R U_R$ is small (less than 20 per cent compared with that of $C_A U_A$). In all cases, the agreement between values of S calculated and those observed is satisfactory. Even for many organic compounds in the liquid state, the heats of evaporation may be calculated, with a fair degree of approximation, as the writer has observed. Thus in these cases the forces of cohesion in the solid and liquid state are of the type discussed in the present chapter.

On the other hand, in some cases a large fraction of the energy of attraction must be due to the presence of intrinsic dipole moments, as for example in the case of H_2O, $C_6H_5NO_2$, and other molecules for which the electric moment μ_e is considerable.

In the case of ionic lattices, such as that of NaCl, the attractive forces are electrostatic, and the corresponding potential energy is given by

$$\phi(\text{attractive}) = -\frac{N}{2} \cdot C_A \cdot \frac{e^2}{r_0},$$

where r_0 is the minimum distance between ions. The problem of cohesion in such lattices has been the subject of a large number of investigations, and the reader who is interested in following it further will find his efforts well repaid.

It should be added that the polarizability α, as well as the proper values of ν_0, necessary for the determination of U_A, may be derived on the basis of wave mechanics from spectral data. However, since such calculations involve the application of the perturbation theory, a dis-

cussion of the methods used as well as of other features of the more recent developments of the theory must be postponed for the following chapter.[33]

COLLATERAL READING

1. The paper by J. E. LENNARD-JONES, *Proc. Phys. Soc.* (*London*), **43**, 461 (1931), presents London's theory in a simpler form than any of the other publications and has been followed to a large extent in the discussion as given in this chapter.
2. For a more detailed discussion of van der Waals forces the reader may consult TAYLOR'S "Treatise," Chapter IV, by K. F. HERZFELD and H. M. SMALLWOOD.
3. C. G. DARWIN, "The New Conceptions of Matter," Macmillan Co., New York, 1931. The reader will find the non-mathematical presentation of difficult concepts in quantum mechanics very stimulating and of invaluable assistance in comprehending the significance of the mathematical arguments as presented in this and other chapters of the present volume.
4. The best references on the topics of polarizability and electric moments are the following:
   (1) P. DEBYE, "Polar Molecules," Chemical Catalog Co., New York, 1929.
   (2) C. P. SMYTH, "Dielectric Constant and Molecular Structure," Chemical Catalog Co., New York, 1931.
5. The equation of state and van der Waals forces are discussed in a comprehensive manner in R. H. FOWLER'S "Statistical Mechanics," Chapter X, Cambridge Univ. Press, Cambridge, England, 2nd edition, 1936.
6. The problem of cohesive energy in ionic lattices is discussed very fully in the following publications:
   (1) M. BORN, "Problems of Atomic Dynamics," Mass. Inst. Technology, Cambridge, Mass., 1926.
   (2) J. SHERMAN, *Chem. Rev.*, **11**, 93 (1932). This gives an excellent summary of the investigations on this topic.
7. Among the more recent papers dealing with the applications of the wave mechanics theory of van der Waals forces the reader will find the following of interest:
   (1) A. MÜLLER, *Proc. Roy. Soc.*, **A154**, 624 (1936). A calculation of the lattice energy of a normal-CH_2 chain molecule in a paraffin crystal.
   (2) J. H. DE BOER, *Trans. Faraday Soc.*, **32**, 10 (1936). Calculation of cohesive energy in resins.
8. PAULING and WILSON, "Quantum Mechanics," Chapter XIV, and SLATER and FRANK, "Introduction to Theoretical Physics," Chapter XXXVI. The reader will find both these discussions extremely well worth studying, but will probably find the first more valuable after he has learned about the perturbation theory and method of variations, and the second, after becoming more familiar with statistical mechanics. (See Chapters IX–XI in this volume for discussion of the first two topics.)

[33] Because of this connection between van der Waals forces and the spectral dispersion curves, these forces have also been designated *dispersion forces*. In an interesting paper in *Trans. Faraday Society*, **28**, 316 (1932), M. Polanyi has discussed the application of this theory to the adsorption of gases.

CHAPTER IX

PERTURBATION THEORY

9.1 Introductory Remarks. The problem of the atom with more than one electron is one which, on the basis of the Bohr theory, has its analog in celestial mechanics, where the problem is that of calculating, for instance, the effect of the gravitational pull of the sun on the motion of the moon around the earth. The field due to the action of the sun is said " to perturb " the motion of the moon, and obviously the effect of the sun must vary with its position in relation to both the moon and earth.

It is the existence of similar perturbations in atomic systems, due to the interaction of the electrons, that makes the consideration of such systems a more complex mathematical problem than that of the hydrogen-like atom. Thus, in the case of the helium atom, the simplest type of many-electron atoms, the total energy is evidently made up of three terms:

(1) the kinetic energy of each electron,
(2) the potential energy due to the *attractive* force between the nucleus and each electron,
(3) the potential energy due to the *repulsive* force between electrons.

If we designate the charge on the helium-like nucleus by Ze, the distance of each electron from the nucleus by r_1 and r_2, respectively, and the interelectronic distance by r_{12}, the total potential energy term is given by

$$V = -\frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}}.$$

The corresponding S. equation is given by

$$\nabla_1^2\phi + \nabla_2^2\phi + \frac{8\pi^2\mu}{h^2}\left(E + \frac{Ze^2}{r_1} + \frac{Ze^2}{r_2} - \frac{e^2}{r_{12}}\right)\phi = 0, \qquad (1)$$

where the subscripts 1 and 2 of the Laplacian operators (in *polar coördinates*) refer to each electron. If the term e^2/r_{12} were omitted, the solution of the S. equation would be identical with that for two *separate*

hydrogen-like atoms of nuclear charge Ze. That is, the solution would be

$$\phi = \phi_{n_1 l_1 m_1}(r_1, \theta_1, \eta_1) \cdot \phi_{n_2 l_2 m_2}(r_2, \theta_2, \eta_2), \tag{2}$$

where r, θ, and η are the coördinate variables of each electron with respect to the nucleus, and ϕ is a characteristic function of the form dealt with previously in discussing the hydrogen-like system. Furthermore, the characteristic energy values would be given by

$$E = E_{n_1} + E_{n_2},$$

where E_{n_1} and E_{n_2} represent the eigenvalues corresponding to the functions ϕ_{n_1} and ϕ_{n_2}, respectively.

It is the presence of the " perturbation " potential energy term e^2/r_{12}, which makes the solution of the S. equation (1) more difficult. A similar perturbation term occurred in the S. equation (8.26) for the interaction of two linear harmonic oscillators, and there the mathematical difficulty was overcome by a transformation of coördinates. By this process it was found possible to separate the variables and thus obtain a product solution of the form indicated in (2). However, it has not been found possible to devise a transformation scheme for equation (1), by which the variables may be separated.

Consequently, it is necessary to use other methods. Of these the most important are the methods based (1) on the application of the perturbation theory and (2) on the application of the calculus of variations. It should be mentioned that the first of these represents the quantum mechanics modification of the method which has been used in calculations of orbits in celestial mechanics, while the second method involves a fundamental principle which has also been applied in classical physics.

The perturbation theory has been described by E. C. Kemble, as " Mostly machinery — a necessary evil!" The derivation of the first- and second-order perturbation terms given in the following sections is practically the same as that given by Condon and Morse.[1] It has the advantage over the derivation given by Schroedinger and other writers that it requires no knowledge of Green's theorem.

In the following sections an attempt has been made to present each of the steps in the derivation as clearly as possible, and for this reason the argument will appear somewhat tedious at best. However, the only excuse that can be given for the whole discussion is that a knowledge of the perturbation theory and of the accompanying concept of matrix elements is absolutely essential for the proper understanding of the manner in which quantum mechanics has solved a vast number of problems presented by atomic and molecular systems.

[1] Condon and Morse, " Quantum Mechanics," Chapter IV.

9.2 Perturbation Theory — First-Order Terms. Let us consider first the non-degenerate case, that is, one in which there is only one eigenfunction corresponding to each discrete energy state. The S. equation for a given system will be of the form

$$\nabla^2 u + \alpha^2(E - V)u = 0, \tag{3}$$

where $\alpha^2 = 8\pi^2\mu/h^2$, and $\nabla^2 u$ may be of the form $\nabla_1^2 u + \nabla_2^2 u + \ldots$, if we are considering an atomic system containing two or more electrons. The potential energy V is of the form

$$V = V_0 + \lambda V_1, \tag{4}$$

where V_0 represents the potential energy in the absence of any perturbation, while λV_1 is the perturbation term. We regard the latter as the product of a term V_1 which depends upon the coördinates and an arbitrary parameter λ, the value of which is small. Thus in the case of the helium-like atom, we could write

$$\frac{V}{Z} = -\left(\frac{e^2}{r_1} + \frac{e^2}{r_2}\right) + \frac{e^2}{Zr_{12}},$$

and regard $1/Z$ as such a variable parameter whose value is small for large values of Z.

The S. equation for the unperturbed system is of the form

$$\nabla^2 u + \alpha^2(E - V_0)u = 0, \tag{5}$$

and we shall assume that solutions of this equation are known. We shall designate these by u_n^0 and the corresponding eigenvalues by E_n^0, so that they satisfy the S. equation

$$\nabla^2 u_n^0 + \alpha^2(E_n^0 - V_0)u_n^0 = 0. \tag{6}$$

Furthermore, we shall assume that these functions u_n^0 are orthonormal. That is,

$$\int \bar{u}_n^0 u_m^0 d\tau = \delta_{nm} = \begin{cases} 0 \text{ for } n \neq m, \\ 1 \text{ for } n = m \end{cases} \tag{7}$$

where $d\tau$ indicates the element of " volume " (which in the case of an atomic system will be of dimension 3ν, if ν represents the total number of electrons), and the integration is carried out over the whole domain in which the coördinate variables have a physical significance. This domain is known as the configuration space and will be designated by this term in the subsequent discussion. The symbol δ_{nm} is used generally in quantum mechanics in order to indicate that the expression may have the value 1 or 0 according as n is equal, or not, to m, and we shall, henceforth, use it in the sense defined in equation (7).

Let us now assume that the functions and energy values which are obtained by solving the S. equation (3), for the perturbed system, may be represented in terms of the functions u_n^0 and E_n^0 by expressions of the form

$$u_n = u_n^0 + \lambda\phi_n + \lambda^2\chi_n, \tag{8}$$

$$E_n = E_n^0 + \lambda\epsilon_n + \lambda^2\eta_n, \tag{9}$$

where ϕ_n and χ_n are functions of the coördinate variables, which represent the first-order and second-order perturbations, respectively, while ϵ_n and η_n represent the corresponding perturbation energy terms of the first and second order, respectively. The parameter λ is identical with that used in equation (4).

If, now, we replace V, u, and E in equation (3) by the expressions in equations (4), (8), and (9), respectively, and carry out the multiplication indicated, the resultant equation is of the form

$$\begin{aligned}
&\frac{1}{\alpha^2}\{\nabla^2 u_n^0 + \lambda\nabla^2\phi_n + \lambda^2\nabla^2\chi_n\} \\
&+ E_n^0 u_n^0 + \lambda E_n^0\phi_n + \lambda^2 E_n^0\chi_n \\
&- V_0 u_n^0 - \lambda V_0\phi_n - \lambda^2 V_0\chi_n \\
&+ \lambda\epsilon_n u_n^0 + \lambda^2\epsilon_n\phi_n \\
&-\lambda V_1 u_n^0 - \lambda^2 V_1\phi_n \\
&+ \lambda^2\eta_n u_n^0 = 0.
\end{aligned} \tag{10}$$

Since λ is an arbitrary parameter, it follows that the coefficient of each power of λ in equation (10) must vanish identically. The coefficient of $\lambda^0 = 1$ is evidently the same as the left-hand side of equation (6) and merely states the obvious conclusion that, for $\lambda = 0$, the solutions of equations (5) and (3) are identical.

The coefficient of λ in equation (10) yields the *inhomogeneous* differential equation

$$\nabla^2\phi_n + \alpha^2(E_n^0 - V_0)\phi_n = \alpha^2(V_1 - \epsilon_n)u_n^0. \tag{11}$$

If the right-hand side of this equation were equal to zero, the equation would be of the *homogeneous* type and identical in form with equation (6). The presence of the expression on the right-hand side of (11) necessitates a special procedure in order to solve the equation.

Since ϕ_n is assumed to be a function of the coördinate variables which is finite and continuous over the whole domain, it may be expressed as a " Fourier's series " in terms of the orthonormalized functions u_k^0. That is, it is possible to develop the function ϕ_n in the form

$$\phi_n = \sum_k B_{nk} u_k^0, \tag{12}$$

where the summation is extended over all integral values of k from 0 to ∞, and any one of the coefficients B_{nk} can be derived by the relation

$$B_{nk} = \int \bar{u}_k^0 \phi_n d\tau,$$

as indicated in the supplementary note, Chapter VI. However, we shall leave the coefficients undetermined for the present. The subscript nk indicates that the coefficient is that of the function u_k^0 and furthermore that it refers to the series development for the function ϕ_n. Thus, for each characteristic solution u_n of the S. equation, there will exist a function ϕ_n which may be represented by an infinite series of terms, each of which is the product of a constant (that is, a magnitude such as B_{nk} which is independent of the coördinate variables) and one of the series of orthonormal functions u_n^0.

Substituting from (12) into (11), the result is

$$\sum_k B_{nk} \nabla^2 u_k^0 + \alpha^2 \sum_k B_{nk}(E_n^0 - V_0) u_k^0 = \alpha^2 (V_1 - \epsilon_n) u_n^0. \tag{13}$$

But, according to equation (6),

$$\nabla^2 u_k^0 - \alpha^2 V_0 u_k^0 = -\alpha^2 E_k^0 u_k^0.$$

Hence equation (13) may be written in the form

$$\sum_k B_{nk}(E_n^0 - E_k^0) u_k^0 = (V_1 - \epsilon_n) u_n^0. \tag{14}$$

Multiplying both sides of (14) by $\bar{u}_n^0$ and integrating over the configuration space, the result is

$$\sum_k B_{nk}(E_n^0 - E_k^0) \int \bar{u}_n^0 u_k^0 d\tau = \int \bar{u}_n^0 V_1 u_n^0 d\tau - \epsilon_n \int \bar{u}_n^0 u_n^0 d\tau. \tag{15}$$

Since ϵ_n is a constant (the perturbation energy of the first order), it may be written outside the sign of integration, whereas V_1 must be kept inside the integration sign because it is a function of the coördinate variables.

Now for $k = n$, $E_n^0 - E_k^0 = 0$, while for $k \neq n$, $\int \bar{u}_n^0 u_k^0 d\tau = 0$. On the other hand, since u_n^0 is a normalized function,

$$\int \bar{u}_n^0 u_n^0 d\tau = 1.$$

Hence, we deduce from (15) the very important result

$$\epsilon_n = \int \bar{u}_n^0 V_1 u_n^0 d\tau. \tag{16a}$$

"Thus," as Condon and Morse remark,[2] "the linear correction to any energy level is simply the average value of the linear term in the potential energy function weighted according to the value of the squared characteristic function at each place. This theorem is the quantum mechanical analog of a similar result in classical mechanics: *The alteration of the energy of a quantum state in the first approximation is equal to the average of the perturbation potential taken over the undisturbed orbit.*"

If u_n^0 is a real function, equation (16*a*) becomes

$$\epsilon_n = \int (u_n^0)^2 V_1 d\tau. \tag{16b}$$

We may now proceed to determine the coefficients B_{nk} in the expansion for ϕ_n, as given by equation (12). If in equation (14) we multiply through by $\bar{u}_j^0$ and integrate over the configuration space, we obtain the relation

$$\sum_k B_{nk}(E_n^0 - E_k^0)\int \bar{u}_j^0 u_k^0 d\tau = \int \bar{u}_j^0 V_1 u_n^0 d\tau - \epsilon_n \int \bar{u}_j^0 u_n^0 d\tau .$$

Because of the orthogonality relation (7), the only non-vanishing term on the left-hand side is that for which $k = j$, while the coefficient of ϵ_n will vanish for $j \neq n$. Hence,

$$B_{nj} = \frac{\int \bar{u}_j^0 V_1 u_n^0 d\tau}{(E_n^0 - E_j^0)}, \tag{17}$$

where $j \neq n$.

To determine the value of B_{nn}, we make use of the requirement that u_n, the perturbed eigenfunction, must also be a normalized function. Hence,

$$\int \bar{u}_n u_n d\tau = 1 = \int \bar{u}_n^0 u_n^0 d\tau + \lambda \int \bar{u}_n^0 \phi_n d\tau + \lambda \int u_n^0 \bar{\phi}_n d\tau$$
$$+ \lambda^2 \int \bar{\phi}_n \phi_n d\tau + \lambda^2 \int \bar{u}_n^0 \chi_n d\tau + \lambda^2 \int u_n^0 \bar{\chi}_n d\tau + \cdots$$

Because of the validity of equation (7) it follows that

$$0 = \lambda\left(\int \bar{u}_n^0 \phi_n d\tau + \int u_n^0 \bar{\phi}_n d\tau\right) + \text{terms in higher powers of } \lambda.$$

Since λ is a variable parameter, each power of λ must vanish identically, and using the series for ϕ_n it follows that

$$\int \bar{u}_n^0 \left(\sum_k B_{nk} u_k^0\right) d\tau + \int u_n^0 \left(\sum_k \bar{B}_{nk} \bar{u}_k^0\right) d\tau = 0.$$

[2] Condon and Morse, "Quantum Mechanics," p. 119.

In this equation, however, each term, except that for which $k = n$, vanishes because of the orthogonality relation (since B_{nk} may be written outside the sign of integration for each term). Consequently,

$$2B_{nn}\int \bar{u}_n^0 u_n d\tau = 0,$$

and hence, $B_{nn} = 0$.

That is, the series for ϕ_n is of the form

$$\phi_n = \sum_k{}' \frac{V_{kn}}{E_n^0 - E_k^0} \cdot u_k^0, \tag{18}$$

where the prime over the summation sign indicates that the value $k = n$ is excluded in the summation, and

$$V_{kn} = \int \bar{u}_k^0 V_1 u_n^0 d\tau. \tag{19}$$

The quantity V_{kn} has been designated, for reasons which will be discussed in a subsequent section, as a "matrix element," and in the following remarks this designation, as well as the symbol V_{kn}, will be used to represent integrals of the type shown in equation (19).

In this connection it is also necessary to point out that, since V_1 is a function of the coördinates, the expression $\bar{u}_n^0 u_k^0 V_1$ is identical with $\bar{u}_n^0 V_1 u_k^0$. However, in view of the fact that in many quantum mechanical problems the matrix elements contain operators which play the same rôle as V_1 (but which are non-commutative), the order of the factors in any one element becomes of prime importance. For this reason it has been considered advisable in the consideration of matrix elements such as V_{kn}, or ϵ_n, to write the expressions for the integrands in the form used in (16*a*) and (19) rather than in that used in (16*b*), even though V_1 may be commuted with the functions $\bar{u}_n^0 u_k^0$ or $\bar{u}_n^0 u_n^0$.

9.3 Second-Order Perturbation Terms. To calculate the *second-order* perturbation terms η_n and χ_n, it is necessary to consider the inhomogeneous differential equation obtained from equation (10) by equating the coefficient of λ^2 to zero. The resulting equation has the form

$$\left(\frac{1}{\alpha^2}\right)\nabla^2\chi_n + (E_n^0 - V_0)\chi_n = -\eta_n u_n^0 - (\epsilon_n - V_1)\phi_n. \tag{20}$$

As in the calculation of the first-order perturbation terms, we assume that χ_n may be represented by a Fourier's series expansion in terms of the orthonormal functions u_n^0. That is, we assume

$$\chi_n = \sum C_{nk} u_k^0, \tag{21}$$

where the coefficients C_{nk} of each term will have to be determined by a procedure analogous to that used for the determination of the coefficients B_{nk} in the series for ϕ_n.

Substituting (21) in (20), the result is

$$\sum_k C_{nk}\left\{\left(\frac{1}{\alpha^2}\right)\nabla^2 u_k^0 - V_0 u_k^0 + E_n^0 u_k^0\right\} = -\eta_n u_n^0 - (\epsilon_n - V_1)\phi_n.$$

But

$$\left(\frac{1}{\alpha^2}\right)\nabla^2 u_k^0 - V_0 u_k^0 = -E_k^0 u_k^0.$$

Hence,

$$\sum_k C_{nk}(E_n^0 - E_k^0)u_k^0 = -\eta_n u_n^0 - (\epsilon_n - V_1)\phi_n. \tag{22}$$

Multiplying both sides of this equation by $\bar{u}_n^0$ and integrating, it follows from considerations similar to those used in deriving the value of ϵ_n that

$$\begin{aligned}
\eta_n &= \int \bar{u}_n^0(V_1 - \epsilon_n)\phi_n d\tau \\
&= \sum_k{}' B_{nk}\int \bar{u}_n^0(V_1 - \epsilon_n)u_k^0 d\tau \\
&= \sum_k{}' \frac{V_{kn}}{(E_n^0 - E_k^0)}\int \bar{u}_n^0(V_1 - \epsilon_n)u_k^0 d\tau \\
&= \sum_k{}' \frac{V_{kn}}{(E_n^0 - E_k^0)}\int \bar{u}_n^0 V_1 u_k^0 d\tau \\
&= \sum_k{}' \frac{V_{kn}\cdot V_{nk}}{(E_n^0 - E_k^0)},
\end{aligned} \tag{23}$$

where V_{nk} is a matrix element defined by equation (19).

The method used for the determination of the second-order perturbation function χ_n is similar to that involved in the derivation of the first-order term ϕ_n.

From equation (22), by using the series for ϕ_n, it follows that

$$\sum_k C_{nk}(E_n^0 - E_k^0)u_k^0 = -\eta_n u_n^0 + \sum_j{}' B_{nj}V_1 u_j^0 - \epsilon_n \sum_j{}' B_{nj}u_j. \tag{24}$$

(The subscript j is used to differentiate the members of the series from the series on the left-hand side.)

Multiplying both sides of equation (24) by $\bar{u}_i^0$ (where i is one of the series of values, 1, 2, . . . j . . . k . . .), and integrating over the configuration space, we obtain the relation

$$C_{ni}(E_n^0 - E_i^0) = \sum_j{}' B_{nj} \int \bar{u}_i^0 V_1 u_j^0 d\tau - \epsilon_n \sum_j{}' B_{nj} \int \bar{u}_i^0 u_j^0 d\tau, \quad (25)$$

since the terms on the left-hand side of (24) which do not involve E_i^0 vanish, while on the right-hand side the term involving η_n vanishes since $i \neq n$.

Now in (25)

$$\int \bar{u}_i^0 u_j^0 d\tau = \delta_{ij} = \begin{cases} 1 \text{ for } j = i \\ 0 \text{ for } j \neq i \end{cases};$$

also for $j = i$

$$\int \bar{u}_i^0 V_1 u_j^0 d\tau = \epsilon_i,$$

while for $j \neq i$

$$\int \bar{u}_i^0 V_1 u_j^0 d\tau = V_{ij}.$$

Substituting these relations as well as that for B_{nj} from equations (17) and (18), it follows that

$$C_{ni} = \frac{(\epsilon_i - \epsilon_n) V_{in}}{(E_n^0 - E_i^0)^2} + \sum_j{}'' \frac{V_{ij} \cdot V_{jn}}{(E_n^0 - E_i^0)(E_n^0 - E_j^0)}, \quad (26)$$

where the two primes over the summation sign indicate that the two values $j = i$ and $j = n$ are *excluded*. It may be shown that the requirement of normalization for the function u_n leads in this case (as in the case of the coefficients B_{nj}) to the conclusion $C_{nn} = 0$.

9.4 Perturbation Terms in Relation to Matrix Elements. It is important for an understanding of the " language " used in discussing problems in quantum mechanics to consider the relationship to matrix elements of the various perturbation terms derived in the previous sections.

A matrix is defined as an *array* of magnitudes in rows and columns. It is merely a table in which certain elements are arranged according to a definite order.

A very simple type of matrix is obtained from a consideration of three linear equations in x, y, z of the form

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= 0 \\ a_2 x + b_2 y + c_2 z &= 0 \\ a_3 x + b_3 y + c_3 z &= 0. \end{aligned}$$

These three equations will yield definite values of x, y, and z if the *determinant*

$$\begin{vmatrix} a_1b_1c_1 \\ a_2b_2c_2 \\ a_3b_3c_3 \end{vmatrix} = 0.$$

The *array* of values

$$\begin{pmatrix} a_1b_1c_1 \\ a_2b_2c_2 \\ a_3b_3c_3 \end{pmatrix},$$

indicated by round brackets, is known as a matrix.

More generally, if we have n equations in the n variables $x_1, x_2, \ldots x_n$, these will be of the form

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \text{---} \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0 \end{array},$$

and the corresponding matrix is

$$\begin{pmatrix} a_{11}a_{12} \ldots a_{1n} \\ a_{21}a_{22} \ldots a_{2n} \\ \text{----------} \\ a_{n1}a_{n2} \ldots a_{nn} \end{pmatrix}.$$

The *element* of the mth row and kth column is indicated as a_{mk}, and it will be observed that the diagonal elements are designated by a_{kk} or a_{mm}.

Now let us consider the integral V_{kn} defined in equation (19), and arrange all the possible values of this integral as a matrix. Since k or n can have any integral value from 1 to ∞, the matrix will consist of a double infinitude of elements, which we can indicate thus:

$$\begin{array}{l} V_{11}\ V_{12}\ V_{13} \ldots V_{1n} \ldots \\ V_{21}\ V_{22}\ V_{23} \ldots V_{2n} \ldots \\ \text{---------------------------------} \\ V_{k1}\ V_{k2}\ V_{k3} \ldots V_{kn} \ldots \\ \text{---------------------------------} \end{array}$$

(Since the determinant will be indicated by enclosing the array in vertical lines, we may omit the brackets when considering the corresponding matrix.)

Now it will be observed that the *first-order perturbation energy terms* are given by the *diagonal elements* of this matrix. That is,

$$\epsilon_n = V_{nn}.$$

The coefficients B_{nk} of the functions u_k^0 in the series expansion

$$\phi_n = \sum_k{}' B_{nk} u_k^0,$$

for the perturbation term of the first order in the eigenfunction, are given, as shown in equation (18), by the relation

$$B_{nk} = \frac{V_{kn}}{(E_n^0 - E_k^0)},$$

where $B_{nn} = 0$. These coefficients evidently correspond to all *the elements of the nth column* with the exception of that element which occurs on the diagonal.

The second-order perturbation energy is given by equation (23), where each term $V_{nk} \cdot V_{kn}$ is obtained by multiplying the k*th element in the* n*th row by the* k*th element in the* n*th column.* The term $(V_{nn})^2$ is excluded from the summation. It will be observed that the method used in deriving the expression for η_n is similar to that used in the multiplication of two determinants.

Since V_1 is a real function of the coördinates, and

$$\begin{aligned} V_{nk} \cdot V_{kn} &= \int \bar{u}_n^0 V_1 u_k^0 d\tau \int \bar{u}_k^0 V_1 u_n^0 d\tau, \\ &= \int \bar{u}_n^0 u_n^0 d\tau \int V_1^2 \bar{u}_k^0 u_k^0 d\tau, \\ &= \int V_1^2 \bar{u}_k^0 u_k^0 d\tau, \end{aligned}$$

it follows that $V_{nk} \cdot V_{kn}$ is a *real* magnitude, and that V_{nk} and V_{kn} are complex conjugates. That is, $| V_{nk} | = | V_{kn} |$. A matrix having the property that elements which are symmetrical with respect to the diagonal form a complex conjugate pair, is known as of the *Hermitian* type. It will be found that all the matrices which occur in quantum mechanics are of this type.

It will also be observed that each term $V_{ij} \cdot V_{jn}$ in the expression given in equation (26) is real and the summation corresponds to the series obtained by multiplying each term in the ith row by the corresponding term in the nth column, omitting those terms which are on the diagonal in each case.

9.5 Perturbation Theory for Degenerate Case.[3] In the previous sections we considered the perturbation theory for the non-degenerate case, that is, the case in which corresponding to each eigenvalue E_n^0

[3] The reader will find this section less difficult after he has studied the discussion of the helium atom problem which will be given in the following chapter.

there exists *only one* eigenfunction u_n^0. In considering the problem of the hydrogen atom it was shown that corresponding to any one eigenvalue E_n^0, there exist n^2 eigenfunctions. This is therefore an illustration of a degenerate state, in which the degeneracy may be removed by the perturbing effect of magnetic or electrostatic fields.

Let us consider the case in which, corresponding to the nth eigenvalue E_n^0, there exist a number ν (which depends upon n) of eigenfunctions

$$u_{n1}^0 \, u_{n2}^0 \cdots u_{n\mu}^0 \cdots$$

In the presence of a perturbing field, the single energy state of quantum number n splits up into ν states for which the energy values may be designated by

$$E_{n\mu} = E_n^0 + \lambda\epsilon_{n\mu}. \tag{27}$$

Since each of the functions $u_{n\mu}^0$ is a solution of the wave equation for the unperturbed state, the wave equation for each of the ν states obtained by removing the degeneracy must reduce, as λ tends to zero, to some linear combination of the functions $u_{n\mu}^0$. Since there are ν functions, it will be possible to form from them ν linearly independent functions of the form

$$u'_{n\mu} = \sum_{\gamma} C_{\mu\gamma} u_{n\gamma}^0, \tag{28}$$

where μ and γ may each assume the values $1, 2, \ldots \nu$. The functions $u'_{n\mu}$ are designated the *zero-order wave functions*, and $C_{\mu\gamma}$ is a coefficient, the value of which varies for each term of the series.

Hence, the first-order approximation to each of the eigenfunctions for the ν states obtained by removing the degeneracy is given by

$$u_{n\mu} = u'_{n\mu} + \lambda\phi_{n\mu}, \tag{29}$$

where $\phi_{n\mu}$ plays a rôle similar to that of ϕ_n in equation (8).

The S. equation for the perturbed states is given by

$$\nabla^2(u'_{n\mu} + \lambda\phi_{n\mu}) + \alpha^2(E_n^0 + \lambda\epsilon_{n\mu} - V_0 - \lambda V_1)(u'_{n\mu} + \lambda\phi_{n\mu}) = 0.$$

It follows that

$$\nabla^2 u'_{n\mu} + \alpha^2(E_n^0 - V_0)u'_{n\mu} = 0; \tag{30}$$

also that

$$\nabla^2\phi_{n\mu} + \alpha^2(E_n^0 - V_0)\phi_{n\mu} = \alpha^2(V_1 - \epsilon_{n\mu})u'_{n\mu}. \tag{31}$$

Evidently equation (30) is satisfied because of the validity of equation (28). To determine ϵ_n we proceed by assuming that $\phi_{n\mu}$ can also be

developed in terms of the unperturbed functions $u^0_{k\beta}$ in the form

$$\phi_{n\mu} = \sum_k \sum_\beta A_{k\beta} u^0_{k\beta}, \tag{32}$$

where the values of the coefficients $A_{k\beta}$ will depend upon the particular values of n and μ as well as on k and β. The notation $u^0_{k\beta}$ refers to any one of the set functions which correspond to the same energy state E^0_k.

Substituting from equation (32) in equation (31), the result is

$$\sum_k \sum_\beta A_{k\beta}(E^0_n - E^0_k) u^0_{k\beta} = (V_1 - \epsilon_n) u'_{n\mu},$$

where ϵ_n is written instead of $\epsilon_{n\mu}$.

Multiplying both sides by $\bar{u}^0_{k\beta}$ and integrating over the configuration space, it is seen that

$$A_{k\beta}(E^0_n - E^0_k) = \int \bar{u}^0_{k\beta}(V_1 - \epsilon_n) u'_{n\mu} d\tau.$$

For $k = n$, this relation becomes

$$0 = \int \bar{u}^0_{n\beta}(V_1 - \epsilon_n) u'_{n\mu} d\tau.$$

By substituting for $u'_{n\mu}$ from equation (28), it follows that

$$\begin{aligned} 0 &= \sum_\gamma C_{\mu\gamma} \left[\int \bar{u}^0_{n\beta} V_1 u^0_{n\gamma} d\tau - \epsilon_n \int \bar{u}^0_{n\beta} u^0_{n\gamma} d\tau \right]; \\ &= \sum_\gamma C_{\mu\gamma} (V^{(n)}_{\beta\gamma} - \epsilon_n \delta^{(n)}_{\beta\gamma}); \end{aligned} \tag{33}$$

where

$$V^{(n)}_{\beta\gamma} = \int \bar{u}^0_{n\beta} V_1 u^0_{n\gamma} d\tau, \tag{34}$$

and

$$\delta^{(n)}_{\beta\gamma} = \int \bar{u}^0_{n\beta} u^0_{n\gamma} d\tau = \begin{cases} 0 \text{ for } \beta \neq \gamma \\ 1 \text{ for } \beta = \gamma \end{cases}. \tag{35}$$

For each value of β there will exist a relation similar to equation (33), and if there are ν eigenfunctions $u^0_{n\beta}$ (or $u^0_{n\gamma}$, $u^0_{n\mu}$, etc.) corresponding to the same energy value E^0_n in the degenerate case, there will be obtained ν linear equations in the ν coefficients $C_{\mu\gamma}$, where $\gamma = 1, 2, \ldots \nu$. Each of these equations will be of the form

$$C_{\mu 1}(V_{\beta 1} - \epsilon_n \delta_{\beta 1}) + C_{\mu 2}(V_{\beta 2} - \epsilon_n \delta_{\beta 2}) + \ldots + C_{\mu\nu}(V_{\beta\nu} - \epsilon_n \delta_{\beta\nu}) = 0.$$

The ν equations thus constitute a system for the determination of the ν constants $C_{\mu 1}$, $C_{\mu 2}$, etc. The condition that such a system has a solution (other than the trivial one that each coefficient vanish identically) is that the determinant of the coefficients of these constants

should vanish. That is, the condition is of the form

$$\left|\begin{array}{lll} V_{11} - \epsilon_n, & V_{12}, & V_{13}, \ldots V_{1\nu} \\ V_{21}, & V_{22} - \epsilon_n, & V_{23}, \ldots V_{2\nu} \\ V_{31}, & V_{32}, & V_{33} - \epsilon_n \cdot V_{3\nu} \\ \hline V_{\nu 1}, & V_{\nu 2}, & V_{\nu 3}, \ldots V_{\nu\nu} - \epsilon_n \end{array}\right| = 0, \qquad (36)$$

where it will be observed that owing to the validity of equation (35) the only coefficients of ϵ_n in equation (33) which are not identically equal to zero are those which are situated on the diagonal of the determinant in (36).

The order of this determinant is evidently the same as the order of degeneracy of the state for which the energy is E_n^0. Consequently, the equation for ϵ_n will be of the form

$$a_n \epsilon_n^\nu + a_{n-1}\epsilon^{\nu-1} + \ldots + a_0 = 0. \qquad (37)$$

That is, it will be an algebraic equation of degree ν, and will possess ν roots.

Equation (36) or (37) is known as the *secular equation.* That the roots are real is evident from the consideration that in the case of real functions $V_{\beta\gamma} = V_{\gamma\beta}$, while in the case of complex conjugate functions it may be demonstrated as in the case of the elements in equation (18) that $V_{\beta\gamma}$ and $V_{\gamma\beta}$ constitute a *complex conjugate pair.*

An illustration of the application of the secular equation to the case $\nu = 2$ will be given in the following chapter in the discussion of the helium atom problem.

Having determined the ν different values of ϵ_n, it becomes possible to obtain *for each of these values* a solution of the set of equations (33) for the ratios of the C's. In other words, for each value $\epsilon_{n\mu}$ a set of values of the ratios of the coefficients $C_{\mu 2} \mid C_{\mu 1}$; $C_{\mu 3} \mid C_{\mu 1}$; $\ldots C_{\mu\nu} \mid C_{\mu 1}$ is obtained, from which the expansion indicated in equation (28) for $u'_{n\mu}$ may be written down.

The procedure for the determination of the coefficients $A_{k\beta}$ in (32) is similar to that used in the non-degenerate case, but more tedious. The relation deduced for $u_{n\mu}$ is of the form

$$u_{n\mu} = \sum_\gamma C_{\mu\gamma} \left\{ u^0_{n\gamma} + \lambda \sum_k \sum_\beta \frac{\int \bar{u}^0_{k\beta} V_1 u^0_{n\gamma} d\tau}{E^0_n - E^0_k} \right\}, \qquad (38)$$

where the summation inside the brackets extends over all states for which $k \neq n$.

9.6 Application to Coupled Linear Oscillators. As mentioned at the beginning of this chapter, the perturbation theory is essentially a

method of transformation by means of which the solution for the perturbed energy state is expressed in terms of the known solution for the unperturbed state and of the expression for the perturbing potential energy function.

It is therefore of interest to consider the solution by the perturbation theory method of a problem for which the solution may also be derived by means of a transformation of coördinate variables. Such a problem is that of the coupled linear harmonic oscillators considered in the previous chapter. The S. equation for the system as shown in that connection is of the form

$$\frac{\partial^2\phi}{\partial q_1^2}+\frac{\partial^2\phi}{\partial q_2^2}+\alpha^2\left(E-\frac{kq_1^2}{2}-\frac{kq_2^2}{2}+\frac{2e^2q_1q_2}{R^3}\right)\phi=0, \qquad (39)$$

where $\alpha^2=8\pi^2\mu/h^2$ and q_1 and q_2 refer to the coördinates of each particle, measured along the same coördinate axis.

It was shown that this equation could be transformed by a change of the variables to the so-called "normal" form, so that the resulting characteristic function could be expressed as the product of two functions (each a function of only one variable). In this case the perturbation term in the expression for the potential energy function is given by $-2e^2q_1q_2/R$, and it was found that this leads to a perturbation term in the expression for the total energy of the form

$$\Delta E=-\frac{h\nu_0e^4}{2k^2R^6},$$

which corresponds to an attractive force between the oscillating particles.

Let us now consider the same problem from the point of view of the perturbation theory.

Using the same symbols as in the problem of the linear harmonic oscillator (see Chapter V), the frequency of each identical oscillator in the unperturbed state is given by

$$\nu_0=\frac{1}{2\pi}\sqrt{\frac{k}{\mu}}.$$

We shall replace q_1 and q_2 in equation (39) by the new variables x_1 and x_2, respectively, defined by the relation

$$\left.\begin{aligned} x=q\sqrt{b}&=q\cdot 2\pi\sqrt{\frac{\mu\nu_0}{h}}\\ &=q\cdot\sqrt{\frac{k}{h\nu_0}}\end{aligned}\right\}. \qquad (40)$$

Consequently, equation (39) becomes

$$\frac{\partial^2\phi}{\partial x_1^2}+\frac{\partial^2\phi}{\partial x_2^2}+\left(\frac{a}{b}-x_1^2-x_2^2+\lambda x_1x_2\right)\phi=0, \tag{41}$$

where $$a=\frac{8\pi^2\mu E}{h^2};$$

and $$\lambda=\frac{16\pi^2\mu e^2}{h^2R^3}\cdot\frac{1}{b^2}=\frac{4e^2}{kR^3}\cdot \tag{42}$$

Thus the perturbation potential energy term is $-\lambda x_1x_2$ where λ is a parameter, the value of which vanishes for the unperturbed state $(R=\infty)$.

The characteristic function for the unperturbed state is

$$u_{nm}^0=\phi_n(x_1)\cdot\phi_m(x_2), \tag{43}$$

and the energy of the system for the quantum states, n and m of each oscillator, is

$$E_{nm}^0=(n+\tfrac{1}{2})h\nu_0+(m+\tfrac{1}{2})h\nu_0=(n+m+1)h\nu_0. \tag{44}$$

As shown in Chapter V, the function $\phi_n(x)$ is real and has the form

$$\phi_n(x)=\epsilon^{-\frac{x^2}{2}}H_n(x),$$

where $H_n(x)$ is the Hermitian polynomial of the nth degree.

Applying equation (16a) for the first-order perturbation energy, which will be designated by $\Delta E^{(1)}$, it follows that

$$\Delta E^{(1)}=-\lambda\int(u_{nm}^0)^2\,x_1x_2dx_1dx_2,$$

where the limits of integration are $\pm\infty$ for each variable. Hence,

$$\Delta E^{(1)}=-\lambda\int\phi_n(x_1)x_1dx_1\int\phi_m(x_1)x_2dx_2. \tag{45}$$

But as was shown in Chapter V, the integrals on the right-hand side of (45) correspond to the average values of x_1 and x_2 respectively, and each of these vanishes. Hence,

$$\Delta E^{(1)}=0.$$

Let us now consider the second-order perturbation energy. Accord-

ing to equation (23) this is given by the relation

$$\Delta E^{(2)} = \lambda^2 \sum{}' \frac{(V_{if})^2}{(E^0_{nm} - E^0_{n'm'})}, \tag{46}$$

where

$$V_{if} = -\int u^0_{nm} x_1 x_2 u^0_{n'm'} dx_1 dx_2, \tag{47}$$

and n' and m' correspond to the values of the quantum numbers n and m for any other state of the system. Thus the total energy of the oscillators in the unperturbed state $n'm'$ is given by

$$E^0_{n'm'} = (n' + m' + 1)h\nu_0.$$

Evidently,

$$V_{if} = -\int \phi_n(x_1)\phi_{n'}(x_1)x_1 dx_1 \int \phi_m(x_2)\phi_{m'}(x_2)x_2 dx_2,$$

and the summation in (46) is extended over all possible values of n' and m', excluding the state $n' = n$, $m' = m$.

Now for the oscillators under consideration, $n = m = 0$. Also, as shown in Chapter V [see equation (5.36) and subsequent discussion], the integral

$$I_{nn'} = \int \phi_n(x)\phi_{n'}(x)x dx$$

vanishes unless $n' - n = \pm 1$. Since $n = 0$, the only possible value of n' is 1, and similarly for m'. Furthermore, since

$$I_{1,0} = \sqrt{\frac{1}{2}},$$

it follows that

$$(V_{if})^2 = \frac{1}{4}.$$

This constitutes the only term in the summation indicated in equation (46). To determine the corresponding values of the denominator $E^0_{00} - E^0_{11}$, it must be observed that in equation (41) both a/b and x are *dimensionless*, and hence in equation (46) $\Delta E^{(2)}$ is expressed as a multiple of a unit of energy. Since $a/b = 2E/h\nu_0$, the unit of energy is $h\nu_0/2$. In terms of these units,

$$E^0_{00} - E^0_{11} = -\frac{2h\nu_0}{\frac{h\nu_0}{2}} = -4.$$

Consequently,

$$\Delta E^{(2)} = -\frac{\lambda^2}{16}\left(\frac{h\nu_0}{2}\right)$$

$$= -\left(\frac{4e^2}{kR^3}\right)^2 \cdot \frac{h\nu_0}{32} = -\frac{h\nu_0 e^4}{2k^2R^6}, \tag{48}$$

which is identical with the result obtained in equation (8.20).

In order to obtain the corresponding perturbation terms in the eigenfunction for the system a much more tedious calculation is necessary, and hence the method used in Chapter VIII is much more convenient in that respect. It should, however, be pointed out that, though $\Delta E^{(1)}$ vanishes, it does not necessarily follow that the corresponding first-order perturbation term in the eigenfunction also vanishes. Thus, if we apply equation (18) and equation (5.36) we obtain the result,

$$\Delta u^{(1)} = -\lambda u_{11}^0 \cdot \frac{\int \phi_0(x_1)\phi_1(x_1)x_1 dx_1 \int \phi_0(x_2)\phi_1(x_2)x_2 dx_2}{E_{00}^0 - E_{11}^0}$$

$$= -\frac{\lambda \cdot \phi_1(x_1)\phi_1(x_2)}{8}$$

$$= -\left(\frac{\lambda}{2}\right)\epsilon^{-\frac{x_1^2+x_2^2}{2}} \cdot x_1 x_2,$$

since $\phi_1(x) = 2x \cdot \epsilon^{-\frac{x^2}{2}}$.

Hence, to a first approximation, the eigenfunction for the system will be given by

$$u = \epsilon^{-\frac{x_1^2+x_2^2}{2}}\left(1 - \frac{\lambda x_1 x_2}{2}\right)$$

$$= u_{00}\left(1 - \frac{\lambda x_1 x_2}{2}\right),$$

which shows that u will be greater or less than u_{00} according as x_1x_2 is positive or negative.

9.7 Interaction Energy of Two Hydrogen Atoms. In Chapter VIII the problem of the interaction of two hydrogen atoms, at such a distance that there is no interchange of electrons, was treated as a problem of three pairs of linear harmonic oscillators, each pair acting along one

of the three orthogonal coördinate-axes. This is the method used by F. London[4] and also by J. E. Lennard-Jones.[5]

On the other hand, R. Eisenschitz and F. London[6] have applied the second-order perturbation theory, using for this purpose the wave functions for the hydrogen atom. The following remarks give a summary of their extremely comprehensive paper.

At relatively small distances of separation two hydrogen atoms may react to form a molecule. The large binding energy arises from the fact that the two electrons, originally associated with separate nuclei, constantly interchange places. This problem was first treated by London and Heitler and is discussed in a subsequent chapter. As the distance of separation between the two atoms increases, the frequency of such interchanges decreases and at sufficiently large distances becomes negligibly low. Under these conditions, however, there exists an attractive energy which is due to dipole-dipole interaction and which, as shown in the previous chapter, accounts for van der Waals forces. The perturbing potential energy term for this interaction is given by the relation

$$V_p = \frac{e^2}{R^3}(x_1x_2 + y_1y_2 - 2z_1z_2). \tag{8.36}$$

In terms of r_1, r_2, and R (see Fig. 50) this may be expressed in the form[7]

$$V_p = \frac{e^2r_1r_2}{R^3}\left[\cos(r_1, r_2) - 3\cos(r_1, R)\cos(r_2, R)\right], \tag{49}$$

where $\cos(r_1, r_2)$ refers to the cosine of the angle between r_1 and r_2, and similarly for $\cos(r_1, R)$ and $\cos(r_2, R)$.

Let $\psi_1(1)$ denote the eigenfunction for the hydrogen atom in which electron 1 occurs and $\phi_1(2)$ the eigenfunction associated with electron (2). Then, the first-order perturbation energy is given in accordance with equation (16*b*) by the relation

$$\Delta E^{(1)} = \int \psi_1^2(1)\phi_1^2(2)V_p d\tau_1 d\tau_2.$$

It may be shown that since the eigenfunctions are spherically symmetrical, the integral in this equation vanishes, so that $\Delta E^{(1)} = 0$.

[4] *Z. physik. Chem.*, **B11**, 222 (1931).

[5] *Proc. Phys. Soc.* (*London*), **43**, 461 (1931).

[6] *Z. Physik*, **60**, 491 (1930).

[7] See Jean's "Treatise on Electricity and Magnetism" for the derivation of this relation.

Let E_n^0 designate the energy of hydrogen atom in state of total quantum number n. Then, in accordance with equation (23), the second-order perturbation energy is given by the relation

$$\Delta E^{(2)} = -\sum_{n}{}' \frac{V_{1n}V_{n1}}{(E_{n'}^0 + E_{n''}^0 - 2E_1^0)}. \tag{50}$$

In this equation, n' and n'' refer to *excited* states from which transitions are permitted to the normal state $n = 1$, and the prime above the summation indicates that we must omit the case $n' = n'' = 1$. These states must therefore be of the nature of p-states ($l = 1$), for which the eigenfunctions are not spherically symmetrical.

The matrix elements are defined by the relation

$$V_{1n} = V_{n1} = \int \psi_1(1)\psi_{n'}(1)V_p\phi_1(2)\phi_{n''}(2)d\tau_1 d\tau_2, \tag{51}$$

in which, in order to simplify the argument, it is assumed that the eigenfunctions are real.

The evaluation of the individual terms in equation (50) is evidently quite involved. By adopting a simple approximation to the series, L. Pauling and E. B. Wilson, Jr.[8] have derived the value

$$\Delta E^{(2)} = -\frac{6e^2a_0^5}{R^6}, \tag{52}$$

while Eisenschitz and London, by carrying out the more laborious computation, obtained a value for $\Delta E^{(2)}$ in which the factor 6 is replaced by 6.47. These conclusions are to be compared with those given in the previous chapter regarding the value of C in equation (8.41).

It is of interest to point out that the physical interpretation of equation (50) is essentially the same as that of equation (46).

Combining equation (49) with equation (51), it is seen that the matrix element V_{1n} involves the product of two integrals which are as follows:

$$\left.\begin{aligned} \mu_{n'} &= e\int r_1\psi_1(1)\psi_{n'}(1)d\tau_1 \\ \text{and} \quad \mu_{n''} &= e\int r_2\phi_1(2)\phi_{n''}(2)d\tau_2 \end{aligned}\right\}. \tag{53}$$

These evidently correspond to dipole moments. Now, as is discussed more fully in Chapter XV, the magnitudes of these moments are related to the intensities of the lines corresponding to the transitions $n' \to 1$ and $n'' \to 1$, respectively. Thus, the value of the second-order

[8] "Quantum Mechanics," pp. 384–385.

perturbation energy may be derived from observations on spectral intensities. It is also evident from these considerations and those stated in the previous chapter that the polarizability α may be calculated from the values of the energy levels and the corresponding eigenfunctions. It is because of the existence of this connection between α and the van der Waals forces that the latter have been designated "dispersion forces." (See footnote at the end of the previous chapter.) The discussion in the following section illustrates even more directly this point of view of the van der Waals forces.

9.8 Application of Perturbation Theory to Stark and Zeeman Effects. In presence of magnetic fields (Zeeman effect) or strong electric fields (Stark effect) spectral lines are resolved into components.[9] Since any spectral line is due to a transition between two energy states, the observed effects must be due to the perturbing effect of the field on the motion of the electron about the nucleus, in each of the two energy states.

A simple illustration of the Stark effect is the problem discussed in Chapter VIII and in the previous section, which dealt with the effect on the energy states and eigenfunctions of a linear harmonic oscillator of the field due to an identical oscillator at a distance.

In the case of the hydrogen atom, the S. equation for the perturbed state in presence of an electric field of strength F applied in the direction of the z-axis is of the form

$$\nabla^2\phi + \frac{8\pi^2\mu}{h^2}\left(E + \frac{e^2}{r} - eFz\right)\phi = 0, \tag{54}$$

where $\lambda V_1 = eFz$ is the perturbing potential energy function.

(In discussing, in Section 8.4, the effect of an electric field on the linear harmonic oscillator, the negative sign was used for the perturbing potential energy function because the electric charge was regarded as a positive magnitude. However, since a field directed in the positive direction of z exerts a force on the electron in the negative direction of z, the positive sign must be used in the present case. It should also be noted that with this convention in signs, the value to be used for the electron charge e should be regarded as a positive magnitude.)

According to equation (16a), the first-order perturbation energy is given by

$$\Delta E^{(1)} = eF\int \bar{\phi}_{nlm} z \phi_{nlm} d\tau. \tag{55}$$

[9] For description of these observations see F. K. Richtmyer's "Introduction to Modern Physics," also H. S. Taylor's "Treatise on Physical Chemistry," Chapter XVI.

Let us consider the case $n = 1$, $l = m = 0$, that is, the normal state of the atom. Then ϕ is a function of r alone, and if $\bar{\phi}\phi$ is multiplied by z and integrated over the whole of the configuration space the integral must vanish. Consequently, $\Delta E^{(1)} = 0$ for the normal state of the hydrogen atom.

In the case of the second-order perturbation term, equation (23) shows that

$$\Delta E^{(2)} = e^2F^2 \sum_k{}' \frac{V_{1k}V_{k1}}{(E_1^0 - E_k^0)}, \tag{56}$$

where

$$V_{1k} = \int \bar{\phi}_{100} z \phi_{klm} d\tau. \tag{57}$$

If $l = m = 0$, then the integral vanishes for the same reason as in the case $n = k = 1$. But if either or both of the eigenfunctions are not spherically symmetrical then V_{1k} has a definite value which may be calculated from the expressions for the eigenfunctions, and from the values of E_k^0.

An interesting application of equation (55) is in the calculation of the polarizability.[10] As pointed out in the previous chapter, the increase in energy of an atom in a uniform electric field of strength F is given by

$$E_p = -\frac{\alpha F^2}{2}. \tag{8.25}$$

But E_p, the polarization energy, must be identical with $\Delta E^{(2)}$, the second-order Stark effect. Hence,

$$\alpha = 2e^2 \sum_k{}' \frac{V_{1k}V_{k1}}{(E_k^0 - E_1^0)}. \tag{58}$$

That is, the polarizability can be calculated, as mentioned in the previous section, from the values of the energy levels and corresponding eigenfunctions.

As shown by E. Schroedinger[11] and subsequent investigators, the relations deduced from equation (54) are in good agreement with the actual observations on the splitting up of the hydrogen lines in an electric field.

In the case of the Zeeman effect the first-order perturbation energy

[10] See references in collateral reading.

[11] *Ann. Physik*, **80**, 437 (1926). See also Condon and Morse, "Quantum Mechanics," pp. 123–129.

term is given by the same relation as was derived in the older theory, that is,

$$\Delta E^{(1)} = \left(\frac{eH}{4\pi\mu c}\right) mh,$$

where m may have any one of the integral values ranging between $-l$ and $+l$, and H is the strength of the magnetic field.

COLLATERAL READING

1. Illustrations of the applications of the perturbation theory are given in (1) PAULING and WILSON, "Quantum Mechanics," Chapter VI, and (2) CONDON and MORSE, "Quantum Mechanics," p. 116 et seq.

2. The method used for deriving equation (58) for the polarizability is that given by LINDSAY and MARGENAU, "Foundations of Physics," pp. 460–469. This section also contains a discussion of the application to the Zeeman effect.

3. The reader will also find it of interest to consult the following references on the perturbation theory:

(1) KEMBLE, E. C., *Phys. Rev. Suppl.*, **1**, 157 (1929).

(2) SLATER and FRANK, "Theoretical Physics," Chapter 32, which develops the relations by use of the secular equation.

4. EISENSCHITZ, R., and LONDON, F., *Z. Physik*, **60**, 491 (1931). The first part of this comprehensive paper is quite difficult to follow, but the reader will find it instructive to study more fully the second part which deals specifically with the problem of the interaction of two hydrogen atoms. This paper is also of interest in connection with the Heitler-London treatment of the problem of the H_2 molecule which is discussed in a subsequent chapter.

CHAPTER X

THE HELIUM ATOM — PERTURBATION METHOD

10.1 Atomic Units. In dealing with the problem of the helium atom and atomic systems in general, it is convenient to use a system of so-called atomic units, which were introduced by D. R. Hartree[1] and have been adopted quite generally by writers on quantum mechanics.[2]

These units are defined as follows:

Unit of length, the radius of the Bohr orbit in the normal state ($n = 1$) of the hydrogen atom, which is designated by a_0, where

$$a_0 = \frac{h^2}{4\pi^2\mu e^2};$$

Unit of charge e, the charge on the electron;

Unit of mass μ, the mass of the electron.

Consistent with these are the following:

Unit of action $h/(2\pi)$, which is usually designated by $\hbar$;

Unit of energy e^2/a_0, which is *twice* the ionization energy of the hydrogen atom with fixed nucleus, that is, 2×13.53 e.v.

Hence in terms of these units

$$r = \sigma a_0, \tag{1}$$

and

$$E = \frac{\lambda \cdot e^2}{a_0} = \frac{\lambda \cdot 4\pi^2\mu e^4}{h^2} \tag{2}$$

$$= \lambda \cdot 2Rch,$$

where σ and λ are dimensionless numbers, and R is the Rydberg constant for infinite mass, that is R_∞.

Let us now consider the radial equation for the hydrogen-like atom which, as shown in equation (7.8), has the form

$$\frac{d^2S}{dr^2} + \frac{2}{r}\cdot\frac{dS}{dr} + \left\{\frac{8\pi^2\mu E}{h^2} + \frac{8\pi^2\mu Ze^2}{h^2 r} - \frac{l(l+1)}{r^2}\right\} S = 0. \tag{3}$$

[1] *Proc. Cambridge Phil. Soc.*, **24**, 89 (1928). An interesting discussion of this topic is given by E. U. Condon and G. H. Shortley in the Appendix of their treatise, "The Theory of Atomic Spectra."

[2] H. Bethe uses these units in his article in "Handbuch der Physik," Vol. XXIV, Part I.

Evidently,

$$\frac{dS}{dr} = \frac{1}{a_0} \cdot \frac{dS}{d\sigma}; \quad \frac{d^2S}{dr^2} = \frac{1}{a_0^2} \cdot \frac{d^2S}{d\sigma^2};$$

$$\frac{8\pi^2\mu E}{h^2} = \frac{\lambda \cdot 32\pi^4\mu^2e^2}{h^4} = \frac{2\lambda}{a_0^2};$$

and

$$\frac{8\pi^2\mu Ze^2}{h^2 r} = \frac{2Z}{a_0^2\sigma}.$$

Hence, equation (3) assumes the form

$$\frac{d^2S}{d\sigma^2} + \frac{2}{\sigma} \cdot \frac{dS}{d\sigma} + \left\{2\lambda + \frac{2Z}{\sigma} - \frac{l(l+1)}{\sigma^2}\right\} S = 0. \tag{4}$$

This has the solution [see equation (7.32)]

$$S_{nl}(\sigma) = \left\{\frac{(n-l-1)!}{2n[(n+l)!]^3} \cdot \left(\frac{2Z}{n}\right)^3\right\}^{\frac{1}{2}} \rho_0^{\,l} \epsilon^{-\frac{\rho_0}{2}} L_{n+l}^{2l+1}(\rho_0), \tag{5}$$

where

$$\rho_0 = \frac{2Z\sigma}{n}. \tag{6}$$

That is,

$$\rho_0 = \frac{2Zr}{na_0}, \tag{7}$$

as in Chapter VII. (The symbol ρ_0 is used to distinguish this from the symbol ρ used subsequently.)

The eigenvalues corresponding to the eigenfunctions $S_{nl}(\sigma)$ are given by the relation

$$\lambda_n = -\frac{Z^2}{2n^2}. \tag{8}$$

That is,

$$E_n = -\frac{4\pi^2\mu e^4}{h^2} \cdot \frac{Z^2}{2n^2} = -\frac{Z^2}{n^2} \cdot Rch.$$

The complete eigenfunction for the system is given, as shown in Chapter VII, by the function

$$\phi_{nlm}(\sigma, \theta, \eta) = S_{nl}(\sigma) \cdot P_l^m(\cos\theta) \cdot \epsilon^{im\eta},$$

which is the solution of the S. equation

$$\nabla^2\phi + \left(2\lambda + \frac{2Z}{\sigma}\right)\phi = 0.$$

For the state $n = 1$, $l = 0$, it follows from equation (7.35) that the normalized eigenfunction is given by

$$\phi_{100} = -\frac{Z^{\frac{3}{2}}}{\sqrt{\pi}} \cdot \epsilon^{-Z\sigma}, \tag{9}$$

and the eigenvalue by

$$\lambda_1 = -\frac{Z^2}{2},$$

that is, by

$$E_1 = -Z^2 \cdot Rch.$$

10.2 The S. Equation for the Helium Atom. For the helium atom, the S. equation in terms of ordinary units [see equation (9.1)] is

$$\nabla_1^2\phi + \nabla_2^2\phi + \frac{8\pi^2\mu}{h^2}\left(E + \frac{Ze^2}{r_1} + \frac{Ze^2}{r_2} - \frac{e^2}{r_{12}}\right)\phi = 0, \tag{10a}$$

where r_1 and r_2 are the distances of the electrons from the nucleus, and r_{12} denotes the interelectronic distance. In terms of Hartree units this equation becomes

$$\nabla_1^2\phi + \nabla_2^2\phi + 2\left(\lambda + \frac{Z}{\sigma_1} + \frac{Z}{\sigma_2} - \frac{1}{\sigma_{12}}\right)\phi = 0, \tag{10b}$$

which may also be written in the form

$$\nabla_1^2\phi + \nabla_2^2\phi + 2\left(\frac{\lambda}{Z} + \frac{1}{\sigma_1} + \frac{1}{\sigma_2} - \frac{1}{Z\sigma_{12}}\right)\phi = 0. \tag{10c}$$

The last equation shows that, for $Z \to \infty$, the perturbation term vanishes, and the solution is then given by the " zero-order " eigenfunction

$$\phi_0 = \phi_{n_1 l_1 m_1}(\sigma_1) \cdot \phi_{n_2 l_2 m_2}(\sigma_2), \tag{11}$$

where σ_1 is used instead of $\sigma_1, \theta_1, \eta_1$, and similarly for σ_2.

The corresponding eigenvalues are evidently given by the relation

$$\begin{aligned}\lambda_0 &= \lambda_1 + \lambda_2 \\ &= -\left(\frac{Z^2}{2n_1^2} + \frac{Z^2}{2n_2^2}\right),\end{aligned} \tag{12}$$

which follows from equation (8).

Evidently the S. equation (10b) or (10c) is unaltered if the coördinates of the two electrons are interchanged. That is, the S. equation is satisfied by either $\phi_1(\sigma_1) \cdot \phi_2(\sigma_2) = u_{12}$ or $\phi_1(\sigma_2)\phi_2(\sigma_1) = u_{21}$. Con-

sequently there must exist a relation of the form

$$u_{21} = c \cdot u_{12},$$

where c is a constant which when multiplied by u_{12} converts it into u_{21}. But if we multiply the function u_{21} by c, we must obtain the function u_{12}. Hence,

$$u_{12} = c^2 \cdot u_{12};\ c = \pm 1,$$

and

$$u_{21} = \pm u_{12}. \tag{13}$$

The zero-order energy corresponding to each of these two functions is the same, as is evident from equation (12). Thus, we have here a *twofold degeneracy*, owing to the fact that the two states defined by u_{21} and u_{12} are indistinguishable when there is no interaction between the electrons. Equation (13) shows that, for an atom with two electrons which are in different quantum states, the eigenfunctions may behave in one of two ways for an interchange of coördinates of the electrons. *The functions either remain unaltered or they change sign.* The former are designated *symmetrical*, and the latter *antisymmetrical*, eigenfunctions. In the case of helium, the symmetrical functions correspond to the para-states, and the antisymmetrical to the ortho-states, as will be shown in a subsequent section.

Only in the normal state ($n_1 = n_2 = 1$, $l_1 = l_2 = 0$) is this degeneracy absent. We shall therefore consider first the solution of the S. equation for this case.

10.3 The Normal State of the Helium Atom. The zero-order eigenfunction for this state is

$$\phi_0 = \phi_1(\sigma_1)\phi_1(\sigma_2) = \frac{Z^3}{\pi} \cdot \epsilon^{-Z(\sigma_1+\sigma_2)}, \tag{14}$$

and the corresponding zero-order energy is evidently

$$\lambda_0 = -2 \cdot \frac{Z^2}{2} = -Z^2.$$

That is, in ordinary units

$$E^{(0)} = -\frac{4\pi^2\mu e^4 \cdot Z^2}{h^2} = -2RchZ^2. \tag{15}$$

Applying the first-order perturbation theory to equation (10*b*), it follows from equation (9.16) that the perturbation energy is given by

$$\Delta\lambda^{(1)} = \int \frac{1}{\sigma_{12}} \phi_1^2(\sigma_1)\phi_1^2(\sigma_2) \cdot 4\pi\sigma_1^2 d\sigma_1 \cdot 4\pi\sigma_2^2 d\sigma_2, \tag{16}$$

where $\phi(\sigma)$ has the form indicated in equation (9).

Let $Z\sigma = \rho$, then the normalized functions are of the form

$$\phi_1^2(\rho_1) = \frac{1}{\pi} \cdot \epsilon^{-2\rho_1}$$

and equation (16) assumes the form

$$\Delta\lambda^{(1)} = Z \int \frac{1}{\rho_{12}} \phi_1^2(\rho_1) \cdot 4\pi\rho_1^2 d\rho_1 \cdot \phi_1^2(\rho_2) \cdot 4\pi\rho_2^2 d\rho_2. \tag{17}$$

The integral in this equation represents, physically, *the energy of interaction of two charge distributions.* Though it could be calculated by expressing ρ_{12} as a function of ρ_1, ρ_2, and θ (the angle between the two radii) by the relation

$$\rho_{12}^2 = \rho_1^2 + \rho_2^2 - 2\rho_1\rho_2 \cos\theta,$$

and then eliminating θ by an integration with respect to this variable, a much simpler method is available, which involves the application of potential theory.[3]

Let us consider a spherically symmetrical charge distribution described by the function $f(\rho)$. The total charge contained in a spherical shell of thickness $d\rho$ at $\rho = \rho_s$ is, therefore, $f(\rho_s) \cdot 4\pi\rho_s^2 d\rho$, and the potential of this shell at any point is given by

$$\left(\frac{1}{\rho}\right) 4\pi\rho_s^2 f(\rho_s) d\rho \quad \text{for} \quad \rho > \rho_s,$$

and by

$$\left(\frac{1}{\rho_s}\right) 4\pi\rho_s^2 f(\rho_s) d\rho = 4\pi\rho_s f(\rho_s) d\rho \quad \text{for} \quad \rho < \rho_s.$$

Hence, the potential at a point $\rho = \rho_s$, due to the total charge distribution represented by

$$f(\rho_1) = \phi^2(\rho_1) = \frac{1}{\pi} \cdot \epsilon^{-2\rho_1},$$

is given by

$$V(\rho_s) = 4\pi \left\{ \frac{1}{\rho_s} \int_0^{\rho_s} \phi^2(\rho_1) \rho_1^2 d\rho_1 + \int_\rho^\infty \phi^2(\rho_1) \rho_1 d\rho_1 \right\}$$

$$= 4 \left\{ \frac{1}{\rho_s} \int_0^{\rho_s} \epsilon^{-2\rho_1} \rho_1^2 d\rho_1 + \int_{\rho_s}^\infty \epsilon^{-2\rho_1} \rho_1 d\rho_1 \right\}.$$

The evaluation of these and similar definite integrals is discussed in

[3] The reader will find an especially interesting discussion of this topic in the paper by A. Unsöld, *Ann. Physik*, **82**, 355 (1926–27). See also Appendix IV, Section 17.

Appendix III. Substituting these values in the last equation, the result obtained is

$$V(\rho_s) = \frac{1}{\rho_s}\left\{1 - \epsilon^{-2\rho_s}(1 + \rho_s)\right\}. \qquad (18)^4$$

Let us now consider the interaction energy of this charge distribution with another charge distribution defined by $\phi^2(\rho_2) = (1/\pi)\epsilon^{-2\rho_2}$. This may be derived from a summation of terms, each of which gives the interaction energy between the charge distribution $\phi^2(\rho_1)$ and the charge located in the shell of thickness $d\rho$ at $\rho = \rho_2$. It is seen that these terms are of the form

$$dU = V(\rho_2) \cdot \phi^2(\rho_2) \cdot 4\pi\rho_2^2 d\rho_2,$$

where ρ_2 takes the place of ρ_s in equation (18).

Hence, the total interaction energy due to the two charge distributions is

$$\begin{aligned} U &= \int_0^\infty V(\rho_2) \cdot \phi^2(\rho_2) \cdot 4\pi\rho_2^2 d\rho_2 \\ &= 4\int_0^\infty \epsilon^{-2\rho_2}\rho_2^2 \left\{\frac{4}{\rho_2}\int_0^{\rho_2} \epsilon^{-2\rho_1}\rho_1^2 d\rho_1 \right. \\ &\quad \left. + 4\int_{\rho_2}^\infty \epsilon^{-2\rho_1}\rho_1 d\rho_1\right\} d\rho_2 \\ &= 4\int_0^\infty \epsilon^{-2\rho}\rho^2 \left\{\frac{1}{\rho} - \epsilon^{-2\rho}\left(1 + \frac{1}{\rho}\right)\right\} d\rho. \qquad (19) \end{aligned}$$

As shown in Appendix III, the evaluation of these integrals leads to the result

$$U = 4\left(\tfrac{1}{4} - \tfrac{1}{32} - \tfrac{1}{16}\right) = \tfrac{5}{8}.$$

Consequently, we deduce from equation (17) the result

$$\Delta\lambda = \tfrac{5}{8}Z,$$

and hence the perturbation energy in ordinary units is given by

$$\begin{aligned} \Delta E &= \tfrac{5}{8}Z \cdot 2Rch \\ &= \tfrac{5}{4}RchZ. \qquad (20) \end{aligned}$$

[4] This equation is the same, except for the difference in units of length, as equation (9) of Unsöld's paper.

Hence, the total energy, corrected for the first-order perturbation, is given [see equation (15)] by

$$E^{(0)} + \Delta E = -2RchZ^2 + \tfrac{5}{4}RchZ.$$

This gives the energy of the helium atom with respect to the *doubly ionized* atom, as represented by the process

$$\mathrm{He} \rightarrow \mathrm{He}^{++} + 2\bar{e}.$$

On the other hand, the energy absorbed in the process

$$\mathrm{He}^{+} \rightarrow \mathrm{He}^{++} + \bar{e}$$

is $RchZ^2$.

Hence, the energy required for the process

$$\mathrm{He} \rightarrow \mathrm{He}^{+} + \bar{e},$$

that is, the ionization energy of the helium atom,[5] is

$$V_i = Rch(Z^2 - \tfrac{5}{4}Z)$$
$$= 13.53 \times 1.5 = 20.30 \text{ volts}.$$

Since the observed value is 24.47 volts, there is a discrepancy of 4.17 volts, that is, of about 17 per cent of the correct value. But it should be noted that without the correction term $\frac{5}{4}Z$, the calculated value for V_i would have been $13.53 \times 4 = 54.1$ volts, so that the correction introduced by use of the first-order perturbation has reduced the calculated value from one which is more than twice the observed value to one which is 17 per cent *lower* than that observed. By calculating the second- and higher-order terms of the perturbation energy, a much better approximation to the correct value could be obtained. However, by the application of the method of variation of parameters (discussed in the following chapter), it has been found possible to avoid the tedious calculations involved in the use of the perturbation theory method, with more satisfactory results.

10.4 Excited States of the Helium Atom. Let us now consider the case in which one of the electrons is in the state $n = 1, l = 0$ (state 1), and the other is in the excited state $n = n, l = 0$ (state n). As pointed out previously we find that in this case it is possible to have two eigenfunctions corresponding to the same zero-order energy value. These may be described thus:

	Electron 1	*Electron* 2	*Eigenfunction*
State:	1	n	$F_{1n} = \phi_1(\rho_1)\phi_n(\rho_2)$
State:	n	1	$F_{n1} = \phi_n(\rho_1)\phi_1(\rho_2)$

[5] Actually the value of R for helium should be used in this calculation. The difference is, however, negligible for the present purpose.

The system is therefore *twofold degenerate*, and we must apply the perturbation theory as developed in Chapter IX for this case. Instead, however, of utilizing directly the results deduced there, it will undoubtedly prove of help to the reader for a proper understanding of the more general discussion given in that chapter, if we apply these arguments to the more specific problem under consideration.

Since the functions F_{1n} and F_{n1} are equally valid solutions of the S. equation for the system, any linear combination of these will also be a solution. We therefore put

$$F = aF_{1n} + bF_{n1}, \tag{21}$$

and write for the solution of the perturbed system

$$\phi = F + \psi,$$

and for the energy of the perturbed system, in Hartree units,

$$\lambda = \lambda_1 + \lambda_2 + \eta. \tag{22}$$

Substituting these in the S. equation (10*b*), we obtain the equation

$$(\nabla_1^2 + \nabla_2^2)(F + \psi) + \left(2\lambda_1 + 2\lambda_2 + 2\eta + \frac{2Z}{\sigma_1} + \frac{2Z}{\sigma_2} - \frac{2}{\sigma_{12}}\right)(F + \psi) = 0.$$

Since F_{1n} and F_{n1} satisfy the S. equation for the unperturbed system,

$$(\nabla_1^2 + \nabla_2^2)F + \left(2\lambda_1 + 2\lambda_2 + \frac{2Z}{\sigma_1} + \frac{2Z}{\sigma_2}\right)F = 0.$$

Neglecting products of the second order, it follows that

$$\nabla^2\psi + \left(2\lambda_1 + 2\lambda_2 + \frac{2Z}{\sigma_1} + \frac{2Z}{\sigma_2}\right)\psi = \left(\frac{2}{\sigma_{12}} - 2\eta\right)F. \tag{23}$$

This equation is inhomogeneous, and in order that it shall have a solution, it is necessary, as shown in the first-order perturbation theory for the non-degenerate case, that the right-hand side of (23) shall be *orthogonal to each of the solutions* F_{1n} and F_{n1} of the corresponding homogeneous equation. Hence, we obtain the *two relations*

$$\int (s - \eta)F \cdot F_{1n} \cdot d\tau = 0, \tag{24}$$

and

$$\int (s - \eta)F \cdot F_{n1} \cdot d\tau = 0, \tag{25}$$

where the integration is carried out over the configuration space defined by the element of " volume,"

$$d\tau = 4\pi\sigma_1^2 d\sigma_1 \cdot 4\pi\sigma_2^2 d\sigma_2, \quad \text{and} \quad s = \frac{1}{\sigma_{12}}.$$

Substituting from (21) it follows that

$$a\int (s-\eta)F_{1n}^2 d\tau + b\int (s-\eta)F_{1n}\cdot F_{n1}d\tau = 0, \tag{26}$$

and

$$a\int (s-\eta)F_{1n}F_{n1}d\tau + b\int (s-\eta)F_{n1}^2 d\tau = 0. \tag{27}$$

Let

$$V_{11} = \int sF_{1n}^2 d\tau = 16\pi^2 \int \frac{1}{\sigma_{12}}\phi_1^2(\sigma_1)\phi_n^2(\sigma_2)\sigma_1^2\sigma_2^2 d\sigma_1 d\sigma_2 \tag{28}$$

$$V_{22} = \int sF_{1n}^2 d\tau = 16\pi^2 \int \frac{1}{\sigma_{12}}\phi_n^2(\sigma_1)\phi_1^2(\sigma_2)\sigma_1^2\sigma_2^2 d\sigma_1 d\sigma_2$$

$$V_{12} = V_{21} = \int sF_{n1}F_{1n}d\tau$$

$$= 16\pi^2 \int \frac{1}{\sigma_{12}}\phi_1(\sigma_1)\phi_n(\sigma_1)\phi_1(\sigma_2)\phi_n(\sigma_2)\sigma_1^2\sigma_2^2 d\sigma_1 d\sigma_2. \tag{29}$$

Now it is evident that

$$\int F_{n1}F_{1n}d\tau = \int \phi_1(\sigma_1)\phi_n(\sigma_1)4\pi\sigma_1^2 d\sigma_1 \int \phi_1(\sigma_2)\phi_n(\sigma_2)4\pi\sigma_2^2 d\sigma_2 = 0.$$

Also that

$$\int F_{n1}^2 d\tau = \int \phi_n^2(\sigma_1)\phi_1^2(\sigma_2)4\pi\sigma_1^2 \cdot 4\pi\sigma_2^2 d\sigma_1 d\sigma_2 = \int F_{1n}^2 d\tau = 1.$$

Furthermore, since the perturbation function $1/\sigma_{12}$ is symmetrical with respect to σ_1 and σ_2, it follows that

$$V_{11} = V_{22}.$$

Hence, we can write equations (26) and (27) in the form

$$a(V_{11} - \eta) + bV_{12} = 0,$$

and

$$aV_{12} + b(V_{11} - \eta) = 0.$$

The solution of this pair of equations is given by solving the " secular " equation

$$\begin{vmatrix} V_{11} - \eta, & V_{12} \\ V_{12}, & V_{11} - \eta \end{vmatrix} = 0, \tag{30}$$

that is,

$$(V_{11} - \eta)^2 - (V_{12})^2 = 0.$$

It follows that

$$\eta = V_{11} \pm V_{12},$$

and

$$a = \pm b.$$

Therefore, the two zero-order eigenfunctions are

$$F_S = a(F_{1n} + F_{n1})$$

and

$$F_A = a(F_{1n} - F_{n1}),$$

where F_S is evidently a *symmetrical* function, since it does not change sign when the coördinates of the two electrons are interchanged, while F_A is *antisymmetrical.*

Since F_S and F_A must each be normalized, it follows that

$$1 = a^2 \int (F_{1n}^2 + F_{n1}^2) d\tau \pm 2a^2 \int F_{1n} F_{n1} d\tau,$$

where the last integral on the right-hand side is equal to zero.

Therefore, $2a^2 = 1$; $a = 1/\sqrt{2}$, and the two zero-order eigenfunctions are

$$F_S = \frac{1}{\sqrt{2}} \left\{ \phi_1(\sigma_1)\phi_n(\sigma_2) + \phi_n(\sigma_1)\phi_1(\sigma_2) \right\}, \tag{31}$$

and

$$F_A = \frac{1}{\sqrt{2}} \left\{ \phi_1(\sigma_1)\phi_n(\sigma_2) - \phi_n(\sigma_1)\phi_1(\sigma_2) \right\}. \tag{32}$$

The corresponding eigenvalues to a first-order approximation are given in Hartree units, by the relations

$$\lambda_S = -\frac{1}{2}\left(Z^2 + \frac{Z^2}{n^2}\right) + V_{11} + V_{12}; \tag{33}$$

$$\lambda_A = -\frac{1}{2}\left(Z^2 + \frac{Z^2}{n^2}\right) + V_{11} - V_{12}; \tag{34}$$

as is evident from equation (12).

The last two relations signify that, as a result of the interaction of the electrons, that state, for which the energy is $-(1 + 1/n^2)RchZ^2$ in

absence of such interaction, is split up into *two states* which differ i energy values.

For the case $n = 2$, as shown in the next section

$$V_{11} = 0.210Z;$$
$$V_{12} = 0.022Z;$$

so that,

$$E_S = \lambda_S \cdot 2Rch = -(1 + \tfrac{1}{4})\, RchZ^2 + 0.464RchZ;$$

and

$$E_A = \lambda_A \cdot 2Rch = -(1 + \tfrac{1}{4})\, RchZ^2 + 0.376RchZ.$$

Thus E_A and E_S represent spectral energy levels which are symmetri cally located with respect to the energy value $E = -(1 + \frac{1}{4})RchZ^2 + 0.420RchZ$, and the *level corresponding to the antisymmetric state* i *lower*, since it is more negative. That is, the binding of the two elec trons to the nucleus is greater for the antisymmetric state. If no repu sive forces existed between the electrons, the binding energy to th nucleus would be $(\frac{5}{4})RchZ^2$. The existence of repulsive forces decrease this energy by $0.420RchZ$, corresponding to the term V_{11}. While th existence of this matrix element is therefore readily interpreted on th basis of classical concepts of interaction of electrical charge distribu tions, the existence of the additional term $\pm V_{12}$, in the expression for th energy, results, as is evident from the previous considerations, from th solution of the S. equation, and is therefore a *non-classical* deductio That this result is a purely quantum mechanical conclusion is of e treme significance, as will be pointed out in another section. Since th occurrence of the term $\pm V_{12}$ is due to the fact that the electrons ma interchange "orbits," integrals of this type, which occur frequentl in quantum mechanics calculations, have been designated as *exchan integrals* in contradistinction to the *Coulomb integrals* of which th term V_{11} is a typical example.

10.5 Calculation of Matrix Elements V_{11} and V_{12}. To determin the values of the matrix elements V_{11} and V_{12}, it is necessary to evaluat the integrals in equations (28) and (29) for specific values of n. As a illustration of such a calculation, we shall consider the case $n =$ $l = 0$, $m = 0$, that is, the state of the electron designated spectr scopically as $2s$. The hydrogen-like eigenfunction for this state i given, in accordance with equation (7.36), by the expression

$$\phi_{200} = \phi_2 = \frac{(\rho - 2) \cdot \epsilon^{-\frac{\rho}{2}}}{\sqrt{N}}, \qquad (35$$

where $\rho = Z\sigma$, and

$$N = \int_0^\infty (\rho - 2)^{\frac{1}{2}} \cdot \epsilon^{-\rho} 4\pi\rho^2 d\rho$$

$$= 4\pi \int_0^\infty (\rho^4 - 4\rho^3 + 4\rho^2)\epsilon^{-\rho} d\rho$$

$$= 32\pi.$$

Thus the problem of evaluating the integral V_{11} is similar to that of the integral in equation (17), and therefore

$$V_{11} = Z \int_{\rho_2=0}^\infty \phi_2^2(\rho_2) 4\pi\rho_2 \left\{ \frac{1}{\rho_2} \int_0^{\rho_2} \phi_1^2(\rho_1) 4\pi\rho_1^2 d\rho_1 + \int_{\rho_2}^\infty \phi_1^2(\rho_1) 4\pi\rho_1 d\rho_1 \right\} d\rho_2;$$

$$= Z \int_0^\infty \frac{\epsilon^{-\rho}(\rho - 2)^2}{32\pi} \cdot \frac{1}{\rho} \{1 - \epsilon^{-2\rho}(1 + \rho)\} 4\pi\rho^2 d\rho;$$

$$= \frac{Z}{8} \int_0^\infty \{\epsilon^{-\rho}(\rho^3 - 4\rho^2 + 4\rho) - \epsilon^{-3\rho}(\rho^4 - 3\rho^3 + 4\rho)\} d\rho;$$

$$= \tfrac{1}{2}Z(\tfrac{1}{2} - \tfrac{13}{162}) = 0.420 RchZ. \tag{36}$$

By analogy with the expression given in the last equation for calculating the integral V_{11}, the integral V_{12} is expressed in the equivalent form

$$V_{12} = Z \int_{\rho_2=0}^\infty \phi_2(\rho_2)\phi_1(\rho_2) 4\pi\rho_2^2 \left\{ \frac{1}{\rho_2} \int_0^{\rho_2} \phi_2(\rho_1)\phi_1(\rho_1) 4\pi\rho_1^2 d\rho_1 + \int_{\rho_2}^\infty \phi_2(\rho_1)\phi_1(\rho_1) 4\pi\rho_1 d\rho_1 \right\} d\rho_2. \tag{37}$$

That is, we replace $\phi_2^2(\rho_2)$ and $\phi_1^2(\rho_1)$ by "mixed" distributions $\phi_2(\rho_2)\phi_1(\rho_2)$ and $\phi_2(\rho_1)\phi_1(\rho_1)$, *as if each electron were continuously alternating between the two quantum states*, so that the effective charge distribution is a geometrical mean, as it were, of the two well-defined distributions. From this point of view the origin of the designation "exchange integral" is readily understood.

Substituting in the integral in (37) from equations (9) and (35), the result is[6]

$$V_{12} = Z\int_0^\infty \frac{\epsilon^{-\frac{3}{2}\rho_2}(\rho_2 - 2)}{4\pi\sqrt{2}} \cdot 4\pi\rho_2^2 \left\{ \frac{1}{\rho_2}\int_0^{\rho_2} \frac{\epsilon^{-\frac{3}{2}\rho_1}(\rho_1 - 2)}{4\pi\sqrt{2}} \cdot 4\pi\rho_1^2 d\rho_1 + \int_{\rho_2}^\infty \frac{\epsilon^{-\frac{3}{2}\rho_1}(\rho_1 - 2)}{4\pi\sqrt{2}} \cdot 4\pi\rho_1 d\rho_1 \right\} d\rho_2 ,$$

$$= Z \cdot \frac{16}{3^6} = 0.0439 RchZ.$$

For the general case in which one electron is in the state $n = 1, l = 0$, and the other in the state $n > 1, l = 0$, the integrals V_{11} and V_{12} may be expressed, as shown by Unsöld, by the relations

$$V_{11} = \frac{1}{2}Z\left(\frac{2}{n^2} - a\right),$$

$$V_{12} = \tfrac{1}{2}Zb$$

in terms of $2Rch$ as a unit of energy. For different values of n, the constants a and b have the values indicated in the following table

n:	2	3	4
a:	0.0802	0.0232	0.00973
b:	0.0439	0.0115	0.00468

The fact that a decreases with increase in n is due to the decrease in Coulomb energy of repulsion between the electrons. The electron wave functions " overlap " less as n is increased. The decrease in the exchange energy V_{12} with increase in n is accounted for on the same basis.

10.6. Exchange Energy. As mentioned already, the existence of terms such as V_{12} is a purely quantum effect which has no classical mechanics analog. The nearest analog to the phenomenon is that of the interaction of linear harmonic oscillators, or that of the electron oscillating between two adjacent potential boxes. As pointed out in Chapter VIII, each of the single energy levels which exist for the unperturbed states of the system (in absence of any coupling action) is split up, as a result of the interaction, into two levels. The difference in energy of any pair of such levels decreases with increase in n. Also corresponding to each of these pairs of levels there exist a symmetric and an antisymmetric eigenfunction.

[6] See footnote 3 of this chapter.

Because of this formal resemblance of the interchange phenomenon in the helium atom to the resonance coupling of two linear harmonic oscillators, the term V_{12} has also been designated as a *resonance energy*. However, it cannot be emphasized too strongly, that any such analogy should not be used to give a physical interpretation of the results deduced in the previous sections. Rather, the occurrence of the term V_{12} must be regarded as a logical consequence of the fact that in the excited state of the helium atom the system is two-fold degenerate.

The difference between the symmetric and antisymmetric states of the excited helium atom may be represented diagrammatically by Fig. 52, which is taken from the excellent discussion of this topic by C. G. Darwin.[7]

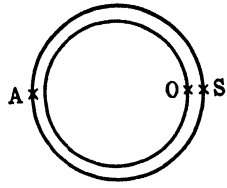


FIG. 52. Symmetry and antisymmetry of electrons. From "The New Conceptions of Matter," by C. G. Darwin.

The two electrons are assumed to be in "orbits" with angular momenta differing by one quantum ($n_1 = 1$, $n_2 = 2$). As Darwin points out, "If one electron is found to be at O, then in the symmetric mode the other is most likely to be at S, whereas in the antisymmetric mode it is most likely to be at A." From this point of view the fact that E_A is greater numerically than E_S is evidently due to the smaller magnitude of the repulsive energy term between the electrons when they are in the antisymmetric state, so that the net attractive energy between each electron and the nucleus is greater.

It is extremely important to realize that the term V_{12} is *not due to the action of any new type of force.* It merely expresses the fact that the two electrons are indistinguishable in their motions. It is thus an indirect result of the validity of the Principle of Indeterminacy.

Instead of the designations V_{kk} and V_{kj} for the Coulomb and exchange integrals, respectively, a number of writers on quantum mechanics have adopted different symbols for the two types of perturbation integrals. Thus Pauling and Wilson use the letters J and K to indicate Coulomb and exchange integrals respectively. For instance,

$$J_s = \int \phi_{10}(1)\phi_{20}(2)\frac{1}{\sigma_{12}}\phi_{10}(1)\phi_{20}(2)d\tau_1 d\tau_2;$$

$$K_s = \int \phi_{10}(1)\phi_{20}(2)\frac{1}{\sigma_{12}}\phi_{10}(2)\phi_{20}(1)d\tau_1 d\tau_2.$$

The subscript s refers to the fact that the eigenfunctions are of the spectral type s.

[7] "The New Conceptions of Matter," p. 205.

As emphasized previously, the integral K arises from the fact that the two electrons are constantly interchanging "orbits." It will also be observed that the existence of this energy term is bound up with the occurrence of a twofold degeneracy, as indicated in equation (29), by the relation

$$V_{12} = V_{21} = \int \frac{1}{\sigma_{12}} F_{1n} F_{n1} d\tau_1 d\tau_2.$$

Furthermore, in the case of the helium atom, this expression has a *positive value.* This result will appear of special significance in the discussion of the hydrogen molecule problem in a subsequent chapter.

Since, in classical mechanics, we define a force as the negative derivative of the energy with respect to distance, in accordance with the relation

$$F = -\frac{\partial E}{\partial r}$$

we could logically consider the exchange energy as due to "exchange forces." However, the form of the expression for V_{12} shows that such a calculation would not only be tedious, but also meaningless.

In fact, we find that in quantum mechanics, the concept "force" is not required. After all, the only magnitude susceptible of measurement is that of the energy involved in the formation of an atom from a nucleus and electrons, or the energy of dissociation of a molecule. Even in the case of van der Waals forces, which were discussed in Chapter VIII, the quantum mechanics calculation led to a value for the interaction energy of the form

$$\Delta E = -\frac{3}{4} \cdot \frac{h\nu_0 \alpha^2}{r^6}.$$

From this it follows as a *mathematical consequence* that

$$-\frac{d(\Delta E)}{dr} \sim \frac{1}{r^7},$$

and if we wish to regard this as defining a force which is a function of r, it is permissible to do so. But actually this deduction does not in any manner help toward a better understanding of the phenomenon.

10.7 Orthohelium and Parhelium. As shown in the previous section, quantum mechanics leads to the conclusion that the spectrum of helium should exhibit two sets of energy levels. These are defined, in

the case $l = 0$ (s-levels) by the relations

$$E_S = -\left(1 + \frac{1}{n^2}\right) RchZ^2 + \left(\frac{2}{n^2} - a\right) RchZ + bRchZ;$$

$$E_A = -\left(1 + \frac{1}{n^2}\right) RchZ^2 + \left(\frac{2}{n^2} - a\right) RchZ - bRchZ.$$

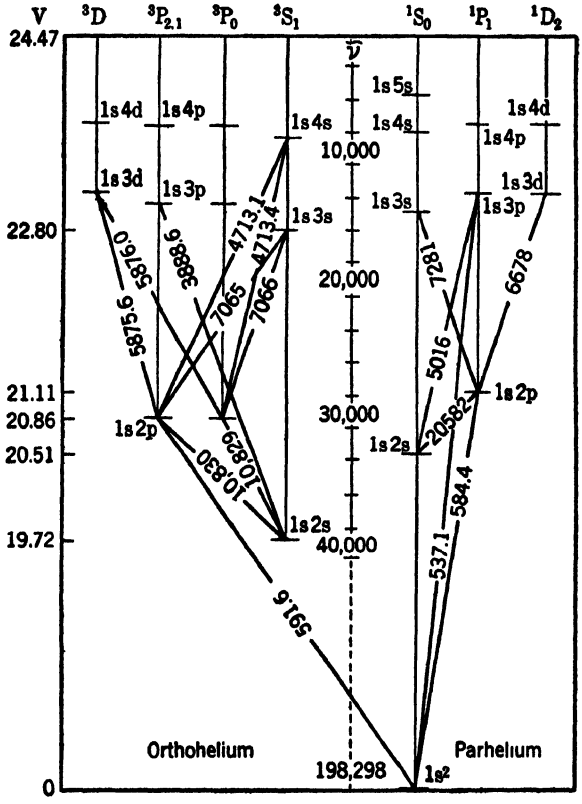


FIG. 53. Energy levels in spectrum of helium.

In terms of *wave numbers* [$\tilde{\nu} = -E/(hc)$] these are given by

$$\tilde{\nu}_{S,A} = RZ^2\left(1 + \frac{1}{n^2}\right) - \frac{2RZ}{n^2} + RZ(a \mp b),$$

where $(a - b)$ refers to $\tilde{\nu}_S$ and $(a + b)$ to $\tilde{\nu}_A$, that is, for any given value of n the term with the smaller numerical value of $\tilde{\nu}$ should correspond to the symmetrical state.

Actually, the spectrum of helium (see Fig. 53) shows two sets of energy levels which were at one time regarded as characteristic of two different atomic species. These are known as the *parhelium* and *orthohelium* series. Each level is designated by the quantum states of the two electrons. Thus, the lowest level is $1s^2$ which indicates that each electron is in the state $n = 1$, $l = 0$ (s-levels), while in the two next higher levels, one electron is in the state $1s$ and the other in the state $2s(n = 2, l = 0)$. It will be observed that, for any given values of n, the level in the parhelium series has the numerically smaller value of the wave number. Consequently, the *parhelium terms must correspond to the symmetrical eigenfunctions* and the *orthohelium terms to the antisymmetrical eigenfunctions.*

The energy diagram also shows that the terms in the orthohelium series are actually triplets, whereas the corresponding terms in the parhelium series are singlets. Thus, there are three levels in the orthohelium series corresponding to the electron configuration $1s2p$ ($n = 1$, $l = 0$ for one electron, and $n = 2$, $l = 1$ for the other electron). Two of these levels (column designated $3P_{2,\,1}$) are so close that it is impossible to indicate them on the diagram as separate levels.

In order to interpret the existence of multiplet levels such as those which occur in the case of helium, it has been found necessary to introduce the concept of *electron spin.* Since this concept and the related generalization, known as Pauli's Exclusion Principle, are of fundamental importance in quantum mechanics, we shall consider these topics in the following section.

10.8 Electron Spin. Pauli's Exclusion Principle. As shown in Chapter VII, the solution of the S. equation for the hydrogen atom leads to a number of eigenfunctions which require for their designation the three quantum numbers, n, l (where $l = n - 1, n - 2, \ldots 0$) and m (where $m = \pm l, \pm(l - 1) \ldots 0$). Since all the eigenfunctions for any given value of n have the same eigenvalue

$$E_n = -\frac{Z^2e^2}{2n^2a_0},$$

we designate this state as *degenerate.* It was also shown that the degree of degeneracy for a given value of n is n^2.

For any given values of l and m, the total angular momentum of the electron in its motion about the nucleus is given by

$$M = \sqrt{l(l+1)} \cdot \frac{h}{2\pi},$$

while the component of this momentum about the Z-axis is

$$M_z = m\frac{h}{2\pi}.$$

In the presence of perturbing fields, such as those produced by the presence of other electrons in the atomic system, or by magnetic or electrostatic fields, the degeneracy may be removed either partly or even completely. Thus in the case of all the atomic systems for which $Z > 1$, the degeneracy with respect to l is removed, so that the states corresponding to different values of l, for the same value of n, have energy values which are no longer identical. These states are designated spectroscopically by the symbols s (for $l = 0$), p (for $l = 1$), d (for $l = 2$), and so forth.

The observations on the effect of magnetic fields on spectral lines (Paschen-Back Effect) show that in a magnetic field the total number of states corresponding to a given value of l is not $2l + 1$ as expected from the previous considerations but *twice this number.* Thus the maximum number of levels for a $1s$, $2s$, or ns electron is 2, in spite of the fact that m can have only the single value 0. Similarly, for an np electron the maximum number of levels is $2(2 \times 1 + 1) = 6$, and for an nd electron, $2(2 \times 2 + 1) = 10$.

Formally, this may be accounted for, as first pointed out by Uhlenbeck and Goudsmit, by the assumption of a spinning electron. The magnitude of the *total* angular momentum of spin is $\sqrt{s(s+1)} \cdot h/(2\pi)$, where s designates the *electron spin quantum number.* The component of this momentum in any given direction is $m_s \cdot (h/2\pi)$, where m_s is a fourth quantum number, the value of which is either $+\frac{1}{2}$ or $-\frac{1}{2}$.

Thus any electron has assigned to it not only the three orbital quantum numbers n, l, m, but also a fourth number known as the magnetic quantum number m_s $(= \pm\frac{1}{2})$ which defines the component of angular momentum with respect to a given axis.

By this theory it is possible to account not only for the observations on the effect of magnetic fields on spectral lines, but also for the existence of multiplet levels as deduced from the spectra of atomic systems containing more than one electron.

For quantum mechanics the significance of these considerations is of importance because of Pauli's Exclusion Principle, which may be enunciated as follows:

In any one atom there cannot exist two electrons having the four quantum numbers (n, l, m, *and* m_s), *respectively, the same in both.*

Thus, if for two electrons, $n_1 = n_2$, $l_1 = l_2$, $m_1 = m_2$, then m_s cannot be identical for both, but must have the value $+\frac{1}{2}$ for one and $-\frac{1}{2}$ for the other electron. That is, two electrons which are identical with respect to each of the three " orbital " quantum numbers, must be antiparallel with respect to the directions of the vectors which correspond to the spin momenta.

Pauli's principle is essentially the real justification for the scheme of electron distribution first suggested by E. C. Stoner and independently by J. D. Main-Smith in 1924. Assuming that the maximum number of electrons of any one type (s, p, d, etc.) which an atomic system may possess is given by the value of $2(2l + 1)$, it is possible to deduce the variation in electron distribution as electrons are added in succession to a nucleus of charge $+Ze$. On this basis an adequate interpretation has been obtained of the periodic arrangement of the elements and of the variation in the characteristics of their spectral terms.

10.9 Multiplet Levels in Spectrum of Helium. From Pauli's principle it follows that in the lowest electronic state (normal state) of the helium atom, in which both electrons are in $1s$ states, the spins must be in *opposite directions*. Since the normal state belongs to the parhelium series, this argument is additional confirmation of the conclusion, stated in a previous section, that the parhelium states correspond to the symmetrical eigenfunctions.

To account for the existence of multiplets it is necessary to introduce into the quantum mechanics formulation of the eigenfunctions the electron spin concept. The procedure used for this purpose is as follows:[8]

It is assumed that corresponding to the electron-spin there exists an eigenfunction $\psi(s)$ which designates the orientation of the axis of spin in a magnetic field. For two electrons there will exist two such functions $\psi(s_1)$ and $\psi(s_2)$, and consequently the *complete* spin function for such a system may be represented by any one of the combinations shown in the attached table. The first two columns give the values of m_s, the third that of $\sum m_s$, and the last column the corresponding combination.

m_{s_1}	m_{s_2}	$\sum m_s$	*Eigenfunction*
$\frac{1}{2}$	$\frac{1}{2}$	1	$\psi(\frac{1}{2})\psi(\frac{1}{2}) = \alpha$
$\frac{1}{2}$	$-\frac{1}{2}$	0	$\psi(\frac{1}{2})\psi(-\frac{1}{2}) = \beta$
$-\frac{1}{2}$	$\frac{1}{2}$	0	$\psi(-\frac{1}{2})\psi(\frac{1}{2}) = \gamma$
$-\frac{1}{2}$	$-\frac{1}{2}$	-1	$\psi(-\frac{1}{2})\psi(-\frac{1}{2}) = \delta$

[8] The following remarks are based upon the discussion in A. Sommerfeld's "Wave Mechanics," pp. 231–233.

The eigenfunctions β and γ evidently represent a degenerate state, since they correspond to interchange of electrons, and we must therefore replace them for the perturbed state of the system by $(\beta + \gamma)/\sqrt{2}$ and $(\beta - \gamma)/\sqrt{2}$. The factor $1/\sqrt{2}$ is introduced for the same reason as it was introduced in equations (31) and (32) for the orbital eigenfunctions F_S and F_A. We thus obtain three spin eigenfunctions which are symmetrical in the electron spins, *viz.*:

$$\alpha, \frac{(\beta + \gamma)}{\sqrt{2}}, \text{ and } \delta,$$

and one which is antisymmetrical, *viz.*: $(\beta - \gamma)/\sqrt{2}$. The first set give rise to triplet terms, for which the values of $\sum m_s$ are 1, 0, and -1, while the last function corresponds to the singlet term $\sum m_s = 0$.

The *complete* eigenfunctions for the system of two electrons are obtained as the *products of orbital and spin functions*. According to equations (31) and (32), the orbital eigenfunctions are

$$F_S = \frac{1}{\sqrt{2}}(F_{1n} + F_{n1}) = u + v,$$

and

$$F_A = \frac{1}{\sqrt{2}}(F_{1n} - F_{n1}) = u - v.$$

Hence, we obtain the following eight combinations of orbital and spin functions which we shall arrange in two sets:

$$\text{(A)}\quad (u+v)\alpha;\ \frac{(u+v)(\beta+\gamma)}{\sqrt{2}};\ (u+v)\delta;\ \frac{(u-v)(\beta-\gamma)}{\sqrt{2}};$$

$$\text{(B)}\quad \frac{(u+v)(\beta-\gamma)}{\sqrt{2}};\ (u-v)\alpha;\ \frac{(u-v)(\beta+\gamma)}{\sqrt{2}};\ (u-v)\delta.$$

Group (A) evidently contains only symmetrical functions, while group (B) consists of the antisymmetrical functions. Now in the energy-level diagram for helium we observe that corresponding to a given value of n and a value of $l > 0$ for the excited state of one of the electrons, there are only *four* levels, three of which belong to the triplet (orthohelium) series and the other to the singlet (parhelium) series. As deduced previously, the latter corresponds to the function which is symmetrical in the orbital functions and antisymmetrical in the spin functions. This evidently means that the parhelium series corresponds

to the eigenfunction $(u + v)\ (\beta - \gamma)/2$ and that the orthohelium series corresponds to the other three members of group (B).

This conclusion may be stated in the following very significant generalization: *The complete solution of the wave equation for any atomic system must involve only that type of eigenfunction which is antisymmetrical,* that is, changes sign when the electrons are interchanged. Evidently, this is the quantum mechanics interpretation of Pauli's Exclusion Principle, and we shall find it extremely important in the consideration of valence problems.

One other deduction which should be mentioned in this connection is the following:

In absence of any interaction between the electron spins and the orbital motions, no transition can occur between a state corresponding to a symmetrical eigenfunction F_S and another state corresponding to an antisymmetrical function F_A. As shown in the energy-level diagram, Fig. 53, there is only one line (λ591.6), and that a faint one, which corresponds to a transition between the ortho- and parhelium series. On the other hand, with atoms of higher nuclear charge, where there is interaction between the electron spin and the orbital motion, lines corresponding to transitions between singlet and triplet levels frequently occur.

COLLATERAL READING

1. UNSÖLD, A., *Ann. Physik*, **82**, 355 (1927). This paper is a classic contribution. It contains a wealth of other material which has not been mentioned in the preceding chapter.

2. On the topic of electron spin in relation to spectral terms the reader should consult H. E. WHITE, "Introduction to Atomic Spectra," and L. A. PAULING and S. GOUDSMIT, "The Structure of Line Spectra."

3. The reader will find a somewhat different method of presentation of spin eigenfunctions in L. PAULING and E. B. WILSON, "Introduction to Quantum Mechanics," Chapter VIII.

CHAPTER XI

THE HELIUM ATOM — VARIATIONAL METHOD

11.1 Minimum Energy Principle. The Variation Theorem. The variational (or Ritz) method has been used extensively in classical physics, especially in the field of dynamics, where its application is at least as old as d'Alembert's principle of virtual displacements. In Chapter IV this principle was discussed as well as the subsequently developed Hamilton's Principle, which was stated in the following form:

For any dynamical path, the time integral of the function $L = L(q_i, \dot{q}_i, t)$, known as the Lagrangian or kinetic potential function, must be an extremum. The function L is defined by the relation

$$L = T - V,$$

where T, the kinetic energy, is a quadratic function of the generalized velocities $\dot{q}_i$ and t, while V, the potential energy, is a function of the generalized coördinates q_i.

Hamilton's Principle can be expressed in the variational form

$$\delta \int_{t_0}^{t_1} L dt = \delta \int_{t_0}^{t_1} (T - V) dt = 0, \tag{4.79}$$

and, as may be shown by the methods of the calculus of variation, the conditions which must be satisfied in order that the integral in the last equation shall be an extremum are given by the f equations, *one* for *each generalized coördinate*, of the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0. \tag{4.36}$$[1]

The last equation is known as *Euler's equation* in the calculus of variation, and also as *Lagrange's equation* in dynamics. It will be observed that in equation (4.79) the *variation applies to a function* L and *not to a coördinate* (as in ordinary calculus), and it is because of this generality that the criterion (4.36) has proved so extremely useful in solving many problems in classical dynamics.

[1] In most cases it will be found that the integral which satisfies the differential equation (4.36) is a minimum. However, the proof of this conclusion would involve much more tedious mathematics.

It was therefore not unreasonable to expect that the method of variations would find equally important application in connection with the problems of quantum mechanics. In fact, the most striking feature about the whole history of the development of theoretical physics during the past two hundred years has been the continuous extension of concepts. The experimental discoveries always seem to demand revolutionary changes in ideas and yet, when in the course of time these observations are arranged in a logical frame, it is found that many of the concepts held previously require only a slight modification, or extension, in order to be able to reconcile them with the new facts.

In the case of quantum mechanics it is readily shown that the *S. equation is essentially the Euler differential equation* which must be satisfied in order that a certain definite integral (which corresponds to the total energy) shall be a minimum.[2] Hence, instead of attempting to solve the S. equation in a particularly difficult case, it may be much more feasible to introduce one or more arbitrary parameters into the expression for the corresponding variational integral. By investigating the effect of variations in the values of these parameters, on the value of the integral, it is then possible to determine a minimum value for the latter, that is, a minimum value for the eigenvalue which corresponds to the given S. equation for a particular energy state.

Let us consider first the S. equation in the *operator* form

$$H\phi = \left(-\frac{1}{\alpha^2}\nabla^2 + V\right)\phi = E\phi, \tag{1}$$

where $\alpha^2 = 8\pi^2\mu/h^2$, and H, known as the *Hamiltonian*, designates the operator in the brackets. Multiplying both sides by $\bar{\phi}$ and integrating over the configuration space, the result is

$$\int \bar{\phi} H\phi d\tau = \int \bar{\phi} E\phi d\tau = E\int \bar{\phi}\phi d\tau,$$

where $d\tau$ is the element of volume in this space. Hence,

$$E = \frac{\int \bar{\phi} H\phi d\tau}{\int \bar{\phi}\phi d\tau}. \tag{2}$$

For normalized functions, the denominator in (2) is equal to unity, and $E = \int \bar{\phi} H\phi d\tau$, where ϕ is an eigenfunction for the given S. equation.

[2] This was pointed out by E. Schroedinger in his first paper on "Quantization as an Eigenvalue Problem," *Ann. Physik*, [4], **79**, 361 (1926).

Suppose, however, that we do not know the exact form of ϕ which should be used to solve equation (1). We may then make a more or less shrewd guess as to the form of the function. Let ψ designate this new function, and let

$$\left.\begin{aligned} I &= \int \bar{\psi} H \psi d\tau \\ \text{and} \qquad N &= \int \bar{\psi}\psi d\tau \end{aligned}\right\}. \tag{3}$$

Then, as shown in the following section, the minimum value of I, which we shall denote by M, is always *greater* (more positive) than the true eigenvalue for the corresponding S. equation. That is,

$$\frac{\int \bar{\psi} H \psi d\tau}{\int \bar{\psi}\psi d\tau} = \frac{M}{N} \gtrless E, \tag{4}$$

or

$$\int \bar{\psi} H \psi d\tau - E \int \bar{\psi}\psi d\tau \gtrless 0. \tag{5}$$

For $\psi = \phi$, the correct eigenfunction, the expression on the left-hand side of the last equation is equal to zero.

To demonstrate the validity of (4) we proceed as follows:[3]

If $E^0 = M/N$ = minimum, then

$$0 = \delta E^0 = \delta\left(\frac{M}{N}\right) = \frac{\delta M}{N} - \frac{M\delta N}{N^2} = \frac{1}{N}(\delta M - E^0 \delta N).$$

Hence,

$$\delta M - E^0 \delta N = 0.$$

That is,

$$\int (\delta\bar{\psi}) H \psi d\tau + \int \bar{\psi} H (\delta\psi) d\tau - E^0 \int (\delta\bar{\psi})\psi d\tau - E^0 \int \bar{\psi}\delta\psi d\tau = 0,$$

or

$$\int (\delta\bar{\psi})\,(H - E^0)\psi d\tau + \int \bar{\psi}(H - E^0)\delta\psi d\tau = 0. \tag{6}$$

Since $\delta\psi$ is arbitrary, let us replace it by $i\delta\psi$.[4] Hence, since $\bar{\psi}\psi = |\,\psi\,|^2$, a real function, we must replace $\delta\bar{\psi}$ by $-i\delta\psi$, in order that the product

[3] A proof of this theorem was first given by C. Eckart, *Phys. Rev.*, **36**, 878 (1930). For the proof given in the text the writer is indebted to Dr. F. Seitz.

[4] This part of the proof is taken from notes on lectures delivered by P. A. M. Dirac at Princeton University in 1931.

$(-i\delta\bar{\psi})$ $(i\delta\psi)$ shall be real. Consequently, (6) becomes

$$-i\int(\delta\bar{\psi})\,(H-E^0)\psi d\tau + i\int\bar{\psi}(H-E^0)\delta\psi d\tau = 0.$$

Dividing both terms by i, and comparing the resulting equation with (6), it follows that the latter can be valid only if *each integral vanishes identically.*[5] Hence,

$$(H-E^0)\psi = 0, \tag{7}$$

which shows that E^0 is an eigenvalue corresponding to the function ψ.

These considerations show that the S. equation is the Euler condition which must be satisfied in order that the integral I in (3) shall be a minimum. The latter may also be expressed in another form, in which it has been used quite frequently by writers on quantum mechanics.

Referring to equation (1), which is merely the S. equation, let us consider the case in which ϕ as well as V are functions of the Cartesian coördinates x, y, z, and ϕ is a real function. (This limitation to real functions and rectangular coördinates simplifies the mathematics but does not detract from the general validity of the conclusions.) Then,

$$\phi H\phi = \phi\left(-\frac{1}{\alpha^2}\nabla^2\phi + V\phi\right) = -\frac{\phi}{\alpha^2}\left(\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}\right)+V\phi^2.$$

But

$$\frac{\phi\partial^2\phi}{\partial x^2} = \frac{d}{dx}\left(\frac{\phi\partial\phi}{\partial x}\right)-\left(\frac{\partial\phi}{\partial x}\right)^2.$$

Hence, designating $dxdydz$ by dv,

$$\int\phi H\phi dv = \frac{1}{\alpha^2}\int\left\{\left(\frac{\partial\phi}{\partial x}\right)^2+\left(\frac{\partial\phi}{\partial y}\right)^2+\left(\frac{\partial\phi}{\partial z}\right)^2\right\}dv$$

$$-\frac{1}{\alpha^2}\left[\frac{\phi\partial\phi}{\partial x}+\frac{\phi\partial\phi}{\partial y}+\frac{\phi\partial\phi}{\partial z}\right]+\int V\phi^2 dv,$$

where the limits of integration are $x = y = z = \pm\infty$. Since ϕ vanishes at the limits of integration the expression in the brackets must be equal to zero. Hence, the integral to be minimized is

$$I = \frac{1}{\alpha^2}\int\left[\left\{\left(\frac{\partial\phi}{\partial x}\right)^2+\left(\frac{\partial\phi}{\partial y}\right)^2+\left(\frac{\partial\phi}{\partial z}\right)^2\right\}+V\phi^2\right]dxdydz,$$

subject to the condition $\int\phi^2 dv = 1$.

[5] An operator β, such that $\bar{\psi}\beta\psi = \psi\beta\bar{\psi}$ is known as *Hermitian* or self-adjoint. Evidently H or $(H-E)$ is an operator of this type.

That is,

$$I - E^0 = \frac{1}{\alpha^2}\int\left[\sum\left(\frac{\partial\phi}{\partial q}\right)^2 - (E^0 - V)\phi^2\right]dv \tag{8}$$

(where q denotes x, y, or z) must be a minimum.

Thus the S. equation is the differential equation which must be satisfied in order that the integral in equation (8) shall be a minimum.

The form of this integral indicates a method by which equation (8) may be derived from the expression for E in the Hamiltonian form

$$E = \frac{1}{2\mu}(p_x^2 + p_y^2 + p_z^2) + V(x, y, z).$$

In this relation let us replace p_x by $(h/2\pi)\,\partial\phi/\partial x$, and similarly for p_y and p_z. Then we obtain the relation

$$E = \frac{1}{\alpha^2}\left\{\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2\right\} + V(x, y, z),$$

in which the expression in the brackets is identical with the first term in the integrand of equation (8).

This may be extended to any system for which the total energy is expressed in the Hamiltonian form, that is, as a function of the f generalized coördinates $q_1 \ldots q_f$, and canonically conjugated momenta $p_1 \ldots p_f$.

Let

$$E = H(p_i, q_i) = T(q_i, p_i) + V(q_i).$$

Under these conditions the element of volume is given by

$$d\tau = \sqrt{A}\cdot dq_1 \ldots dq_f,$$

where $\sqrt{A}$ is known as the " discriminant,"[6] and therefore the integral to be minimized is

$$I = \int\left\{\frac{h^2}{4\pi^2}\cdot T\left(q_i, \frac{\partial\phi}{\partial q_i}\right) + V\phi^2\right\}d\tau, \tag{9}$$

with the condition that

$$\int\phi^2 d\tau = \int\phi^2\sqrt{A}\cdot dq_i \ldots dq_f = 1,$$

where the integration is carried out over the whole configuration space.

[6] This was discussed in Chapter VI in connection with the form of the Laplacian operator. The discriminant is thus identical in the case of three coördinates, with the magnitude $\sqrt{a_1a_2a_3}$ used in equation (6.3).

Now in any treatise on the calculus of variation it is shown that, given a definite integral of the form

$$I = \int F\left(q_i, \phi, \frac{\partial \phi}{\partial q_i}\right) d\tau,$$

the condition to be satisfied in order that this integral shall be a minimum is given by the Euler equation

$$\sum_i \frac{\partial}{\partial q_i}\left(\frac{\partial F}{\partial \phi_j}\right) - \frac{\partial F}{\partial \phi} = 0, \tag{10}$$

where $\phi_j = \partial\phi/\partial q_j$, and i and j are any pair of the numbers $1, 2 \ldots f$.

If we apply this condition to the integral in (9) we obtain the S. equation

$$\frac{h^2}{4\pi^2}\frac{1}{\sqrt{A}}\sum_i \frac{\partial}{\partial q_i}\left(\sqrt{A}\sum_k A_{ik}\frac{\partial \phi}{\partial q_k}\right) + (E^0 - V)\phi = 0, \tag{11}$$

where E^0 is the minimum value of I in (9) for a given form of ϕ, and A_{ik} is a coefficient of the term involving $\partial^2\phi/(\partial q_i \partial q_k)$.

Thus equation (10) leads to the expression for the Laplacian operator in the S. equation[7]

$$\frac{h^2}{8\pi^2\mu}\nabla^2\phi + (E - V)\phi = 0.$$

As an illustration of the application of these equations, let us consider the problem of the linear harmonic oscillator, which was discussed in Chapter V.

In terms of q, the displacement, and p, the corresponding momentum,

$$E = H(p, q) = \frac{p^2}{2\mu} + \frac{kq^2}{2}.$$

Hence, the variational function is given by

$$I = \int_{-\infty}^{\infty}\left\{\frac{h^2}{8\pi^2\mu}\left(\frac{\partial \phi}{\partial q}\right)^2 + \frac{kq^2\phi^2}{2}\right\} dq = \text{minimum} = E,$$

with the condition $\int_{-\infty}^{\infty}\phi^2 dq = 1$.

[7] See also E. U. Condon and P. M. Morse, "Quantum Mechanics," Chapter I, and E. C. Kemble, *Phys. Rev. Suppl.*, **1**, 157 (1929), for more detailed discussion of this topic.

We can combine these two relations in the form, analogous to equation (11),

$$\delta(I - E) = \delta \int_{-\infty}^{\infty} \left\{ \frac{h^2}{8\pi^2\mu} \left(\frac{d\phi}{dq} \right)^2 + \frac{kq^2\phi^2}{2} - E\phi^2 \right\} dq = 0.$$

Denoting the expression in the integrand by $F(q, \phi, d\phi/dq)$, it is evident, if we denote $d\phi/dq$ by ϕ_q, that

$$\frac{\partial F}{\partial \phi_q} = \frac{2h^2}{8\pi^2\mu} \cdot \frac{d\phi}{dq}.$$

Hence,

$$\frac{d}{dq}\left(\frac{\partial F}{\partial \phi_q} \right) = \frac{2h^2}{8\pi^2\mu} \cdot \frac{d^2\phi}{dq^2};$$

$$\frac{\partial F}{\partial \phi} = 2\left(\frac{kq^2}{2} - E \right)\phi;$$

and

$$\frac{d}{dq}\left(\frac{\partial F}{\partial \phi_q} \right) - \frac{\partial F}{\partial \phi} = \frac{2h^2}{8\pi^2\mu} \cdot \frac{d^2\phi}{dq^2} + 2\left(E - \frac{kq^2}{2} \right)\phi$$

$$= 0 \text{ if } (I - E) \text{ is to be a minimum.}$$

The right-hand expression in the last equation is obviously the same as the S. equation (*5*.5) for the linear harmonic oscillator.

11.2 The Energy for the Normal State of the Helium Atom. In the previous chapter the energy of the normal state was calculated by use of the relation for the first-order perturbation energy. The value derived in this manner for the ionization potential of He was found to be 4.17 volts too low. It will be shown in this section that by the use of the variational method it is possible to deduce a value which is much nearer to the spectroscopically correct value.

As shown in equation (*10*.10*b*) the S. equation for the system, in which the eigenvalue is expressed in terms of Hartree units, has the form

$$\nabla_1^2\phi + \nabla_2^2\phi + 2\left(\lambda + \frac{Z}{\sigma_1} + \frac{Z}{\sigma_2} - \frac{1}{\sigma_{12}} \right)\phi = 0. \tag{12}$$

Let us assume a *normalized* solution of the form

$$\phi_0 = \phi(\sigma_1) \cdot \phi(\sigma_2) = \frac{\beta^3 Z^3}{\pi} \cdot \epsilon^{-\beta Z(\sigma_1 + \sigma_2)}, \tag{13}$$

where β is a parameter which has to be determined from the condition

$$\frac{\partial \lambda}{\partial \beta} = 0. \tag{14}$$

Equation (13) means, physically, that we choose a solution in which the nuclear charge Ze is replaced by the effective charge βZe.

For the state $n = 1$ of a hydrogen-like atom of charge βZe, the S. equation has the form, as shown in Section 10.1,

$$\nabla^2 \phi + 2\left(\lambda + \frac{\beta Z}{\sigma}\right)\phi = 0. \tag{15}$$

This leads to the eigenvalue

$$\lambda = -\frac{\beta^2 Z^2}{2},$$

and the eigenfunction

$$\phi(\sigma) = -\frac{(\beta Z)^{\frac{3}{2}}}{\sqrt{\pi}} \epsilon^{-\beta Z \sigma}. \tag{16}$$

Thus, the suggestion of a solution of the form indicated in equation (13) may be interpreted thus:

We regard the repulsive interaction of the electrons as equivalent in effect to a decreased force of attraction between each electron and the nucleus. That is, instead of regarding each electron as acted upon by a nuclear charge of magnitude Ze, we replace the latter by an effective charge of magnitude $\beta \cdot Ze$, where $\beta < 1$. From this point of view we may also consider that each electron screens the other electron to some extent and hence we may designate $(1 - \beta)$ as the "screening constant."[8]

The application of the principle of minimum energy leads to the relation

$$\int \phi_0 H \phi_0 d\tau = \text{minimum} = \lambda^0 \int \phi_0^2 d\tau. \tag{17}$$

For λ^0 expressed in Hartree units, the operator H is defined, in consequence of equation (12), by the relation (see Supplementary Note)

$$H = -\tfrac{1}{2}\left(\nabla_1^2 + \nabla_2^2\right) - \left(\frac{Z}{\sigma_1} + \frac{Z}{\sigma_2} - \frac{1}{\sigma_{12}}\right). \tag{18}$$

[8] C. Eckart, *Phys. Rev.*, **36**, 878 (1930).

Hence,

$$H\phi_0 = -\tfrac{1}{2}[\phi(\sigma_2)\nabla^2\phi(\sigma_1) + \phi(\sigma_1)\nabla^2\phi(\sigma_2)]$$
$$-\left[Z\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}\right)-\frac{1}{\sigma_{12}}\right]\phi(\sigma_1)\phi(\sigma_2).$$

But from equation (15) it is seen that

$$\nabla^2\phi(\sigma_1) = \left(\beta^2Z^2 - \frac{2\beta Z}{\sigma_1}\right)\phi(\sigma_1),$$

and a similar relation applies to $\phi(\sigma_2)$.

Therefore,

$$\phi_0 H\phi_0 = \left\{\phi(\sigma_1)\phi(\sigma_2)\right\}^2\left\{-\beta^2Z^2 + \beta Z\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}\right)\right.$$
$$\left.-Z\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}\right)+\frac{1}{\sigma_{12}}\right\}, \tag{19}$$

and

$$\int \phi_0 H\phi_0 d\tau = J_1 + J_2 + J_3,$$

where

$$J_1 = -\beta^2Z^2\int\phi^2(\sigma_1)d\tau_1\int\phi^2(\sigma_2)d\tau_2,$$

$$J_2 = Z(\beta-1)\int\left(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}\right)\phi^2(\sigma_1)\phi^2(\sigma_2)d\tau_1 d\tau_2,$$

and

$$J_3 = \int\frac{1}{\sigma_{12}}\phi^2(\sigma_1)\phi^2(\sigma_2)d\tau_1 d\tau_2.$$

Since the functions are normalized,

$$J_1 = -\beta^2Z^2.$$

Since the integral J_2 must be symmetrical in σ_1 and σ_2,

$$J_2 = 2Z(\beta-1)\int\frac{\phi^2(\sigma_1)}{\sigma_1}\phi^2(\sigma_2)d\tau_1 d\tau_2$$
$$= 32Z(\beta-1)\cdot\beta^6Z^6\int\epsilon^{-2\beta Z\sigma_1}\sigma_1\cdot d\sigma_1\int\epsilon^{-2\beta Z\sigma_2}\sigma_2^2 d\sigma_2,$$
$$= 2\beta(\beta-1)Z^2.$$

Also, from equations (*10*.16) and (*10*.20) we obtain the relation

$$J_3 = \tfrac{5}{8}\beta Z.$$

Hence,

$$\lambda^0 \int \phi_0^2 d\tau = \lambda_0 = J_1 + J_2 + J_3$$

$$= Z^2 \left\{ -\beta^2 + 2\beta(\beta - 1) + \frac{5\beta}{8Z} \right\}.$$

For $Z = 2$,

$$\lambda^0 = Z^2 \left\{ -\beta^2 + 2\beta(\beta - 1) + \frac{5\beta}{16} \right\},$$

and

$$\lambda^0 = Z^2 \left(\beta^2 - \frac{27\beta}{16} \right). \tag{20}$$

Equation (20) gives λ^0, the value of the energy in Hartree units, as a function of the parameter β. This will be a minimum for the value of β determined by the condition

$$\frac{d\lambda^0}{d\beta} = Z^2 \left(2\beta - \frac{27}{16} \right) = 0.$$

That is,

$$\beta = \tfrac{27}{32},$$

from which we obtain for the value of E^0, the energy in ordinary units, the relation

$$E^0 = 2RchZ^2\beta \left(\beta - \frac{27}{16} \right)$$

$$= -\frac{729RchZ^2}{512}. \tag{21}$$

This represents the energy of formation of He from the two electrons and a nucleus of charge Ze. Hence, the ionization potential of the helium atom is given by

$$V_i = \tfrac{217}{512} \times 4 \times 13.53 \text{ volts}$$

$$= 22.94 \text{ volts}.$$

It will be observed that this value is 1.53 volts less than the observed value (24.47 volts), whereas the first-order perturbation relation leads to a discrepancy of 4.17 volts.

The value $\beta = \frac{27}{32}$ means that the effective nuclear charge is $\frac{27}{16}$ in units of e (the charge on the electron), and that consequently the interaction of the electrons is equivalent to a decrease of $\frac{5}{16}$ units in the nuclear charge.

Instead of the form assumed for the function ϕ_0 in the preceding discussion it is obviously possible to find other forms which lead to minima that approach the true value more closely. A number of theoretical physicists have attacked this problem by more involved mathematical methods,[9] and of these the most successful results have been obtained by E. A. Hylleraas.[10]

The method used by him involves the choice of a system of coördinates which has also proved suitable for other problems in quantum mechanics. For this reason a summary is given in the following section of his technic.

11.3 Variational Method of Hylleraas. In the following section we shall replace σ_1, σ_2, and σ_{12} by r_1, r_2, and r_{12}, respectively. In terms of these coördinates the element of volume is derived as follows. Let r_1, η, ζ denote the polar coördinates of the first electron, and r_2, θ, χ, those of the second referred to the vector r_1 as polar axis. The angle between r_1 and r_2 is given by θ, and hence

$$r_{12}^2 = r_1^2 + r_2^2 - 2r_1r_2 \cos\theta,$$

so that

$$r_{12}dr_{12} = r_1r_2 \sin\theta d\theta.$$

The element of volume is given by

$$d\tau' = r_2^2 \sin\theta\, dr_2 d\theta d\chi \cdot r_1^2 \sin\eta dr_1 d\eta d\zeta$$
$$= r_1r_2r_{12}dr_1dr_2dr_{12}d\chi \cdot \sin\eta d\eta d\zeta.$$

For spherically symmetrical distributions, we may replace $d\tau'$ by the element of volume

$$d\tau = r_1r_2r_{12}dr_1dr_2dr_{12}\int_0^{\pi} \sin\eta d\eta \int_0^{2\pi} d\chi \int_0^{2\pi} d\zeta$$
$$= 8\pi^2 \cdot r_1r_2r_{12}dr_1dr_2dr_{12}. \qquad (22)$$

[9] A list of the variational functions used and of the values of E calculated by each of these methods is given by L. Pauling and E. B. Wilson, "Introduction to Quantum Mechanics," p. 224.

[10] E. A. Hylleraas, *Z. Physik*, **54**, 347 (1929); see also *ibid.*, **48**, 469 (1928); **65**, 209 (1930).

Now let us introduce the so-called " elliptical " coördinates

$$s = r_1 + r_2, \tag{23}$$

$$t = r_1 - r_2, \tag{24}$$

and let

$$u = r_{12}. \tag{25}$$

Since $4r_1r_2 = s^2 - t^2$,

$$d\tau = 2\pi^2(s^2 - t^2)u \cdot du \cdot dr_1dr_2. \tag{26}$$

In order to express the element of area dr_1dr_2 in terms of $ds \cdot dt$ we make use of the following relation between these elements (which is derived in any treatise on integral calculus):[11]

$$\left(\frac{\partial t}{\partial r_1} \cdot \frac{\partial s}{\partial r_2} - \frac{\partial t}{\partial r_2} \cdot \frac{\partial s}{\partial r_1}\right) dr_1dr_2 = ds \cdot dt. \tag{27}$$

From the functional relations between s, t, r_1, and r_2 in equations (23) and (24) it is evident that

$$\frac{\partial s}{\partial r_1} = 1; \quad \frac{\partial t}{\partial r_1} = 1; \quad \frac{\partial s}{\partial r_2} = 1; \quad \frac{\partial t}{\partial r_2} = -1.$$

Consequently (27) becomes

$$2dr_1dr_2 = ds \cdot dt,$$

and (26) becomes

$$d\tau = \pi^2(s^2 - t^2)u \cdot du \cdot ds \cdot dt. \tag{28}$$

The limits of integration are evidently

$$-u \leqslant t \leqslant u,$$
$$0 \leqslant u \leqslant s \leqslant \infty.$$

Let $\phi(s, t, u)$ denote some function of the coördinates. This will be a solution of the S. equation (12) if the following condition is satisfied.

$$I = \frac{1}{N}\int \phi H \phi d\tau = \text{minimum} = \lambda,$$

where

$$N = \int \phi^2 d\tau.$$

The Hamiltonian operator has the form

$$H = -\tfrac{1}{2}(\nabla_1^2 + \nabla_2^2) + V,$$

[11] The expression in the parentheses is known as a Jacobian.

where

$$V = -\left(\frac{Z}{r_1}+\frac{Z}{r_2}-\frac{1}{r_{12}}\right). \tag{29}$$

By application of Green's theorem, the expression

$$\phi H\phi = -\tfrac{1}{2}\phi(\nabla_1^2+\nabla_2^2)\phi + V\phi^2 \tag{30}$$

can be transformed into a form more suitable for integration. This theorem (see Appendix IV) states that, if U is a function of the rectangular coördinates x, y, z, then

$$\int\int\int U\nabla^2 U dxdydz + \int\int\int\left\{\left(\frac{\partial U}{\partial x}\right)^2+\left(\frac{\partial U}{\partial y}\right)^2+\left(\frac{\partial U}{\partial z}\right)^2\right\}dxdydz + \int\int U\frac{\partial U}{\partial n}ds = 0, \tag{31}$$

where $\partial U/\partial n$ is the rate of change of U along the normal at any point on a surface described in the configuration space, and ds is an element of area on this surface. If the integration indicated in the first two integrals is carried out over the whole configuration space (that is, with the limits $x = y = z = \pm\infty$), the values of U and $\partial U/\partial n$ on the surface described at these limits must vanish, if U is a solution of the S. equation. Also the integrand in the second term of this equation is evidently equal to $(\partial U/\partial n)^2$, that is, to the square of the gradient along the diagonal of the element of volume $dxdydz$. Hence, we can express equation (31) in the form

$$\int U\nabla^2 U dxdydz = -\int (\mathrm{grad}U)^2 dxdydz, \tag{32}$$

where the limits of integration are $x = y = z = \pm\infty$.

In the case of the function $\phi(r_1, r_2, r_{12})$,

$$r_1^2 = x_1^2 + y_1^2 + z_1^2,$$
$$r_2^2 = x_2^2 + y_2^2 + z_2^2,$$

and $$r_{12}^2 = (x_1-x_2)^2 + (y_1-y_2)^2 + (z_1-z_2)^2,$$

where $x_1y_1z_1$ are the coördinates of the first electron and $x_2y_2z_2$ refer to the second electron.

Now

$$\frac{\partial\phi}{\partial x_1} = \frac{\partial\phi}{\partial r_1}\cdot\frac{\partial r_1}{\partial x_1}+\frac{\partial\phi}{\partial r_{12}}\cdot\frac{\partial r_{12}}{\partial x_1}$$

$$= \frac{\partial\phi}{\partial r_1}\cdot\frac{x_1}{r_1}+\frac{\partial\phi}{\partial r_{12}}\cdot\frac{x_1-x_2}{r_{12}}.$$

Hence, for the first electron

$$(\mathrm{grad}_1\phi)^2 = \left(\frac{\partial\phi}{\partial x_1}\right)^2 + \left(\frac{\partial\phi}{\partial y_1}\right)^2 + \left(\frac{\partial\phi}{\partial z_1}\right)^2$$

$$= \left(\frac{\partial\phi}{\partial r_1}\right)^2 + \left(\frac{\partial\phi}{\partial r_{12}}\right)^2 + \frac{\partial\phi}{\partial r_1}\cdot\frac{\partial\phi}{\partial r_{12}}\left\{\frac{2x_1(x_1 - x_2)}{r_1 r_{12}}\right.$$

$$\left. + \frac{2y_1(y_1 - y_2) + 2z_1(z_1 - z_2)}{r_1 r_{12}}\right\}$$

$$= \left(\frac{\partial\phi}{\partial r_1}\right)^2 + \left(\frac{\partial\phi}{\partial r_{12}}\right)^2 + \frac{\partial\phi}{\partial r_1}\cdot\frac{\partial\phi}{\partial r_{12}}\cdot\frac{r_1^2 + r_{12}^2 - r_2^2}{r_1 r_{12}},$$

and a similar expression may be derived for $(\mathrm{grad}_2\phi)^2$, which refers to the second electron.

In terms of the variables s, t, u,

$$\frac{\partial\phi}{\partial r_1} = \frac{\partial\phi}{\partial s}\cdot\frac{\partial s}{\partial r_1} + \frac{\partial\phi}{\partial t}\cdot\frac{\partial t}{\partial r_1} = \frac{\partial\phi}{\partial s} + \frac{\partial\phi}{\partial t};$$

$$\frac{\partial\phi}{\partial r_2} = \frac{\partial\phi}{\partial s} - \frac{\partial\phi}{\partial t};$$

$$\frac{\partial\phi}{\partial r_{12}} = \frac{\partial\phi}{\partial u}.$$

Hence,

$$(\mathrm{grad}_1\phi)^2 + (\mathrm{grad}_2\phi)^2 = 2\left(\frac{\partial\phi}{\partial s}\right)^2 + 2\left(\frac{\partial\phi}{\partial t}\right)^2 + 2\left(\frac{\partial\phi}{\partial u}\right)^2$$

$$+ \frac{4}{u(s^2 - t^2)}\cdot\frac{\partial\phi}{\partial u}\left[s(u^2 - t^2)\frac{\partial\phi}{\partial s} + t(s^2 - u^2)\frac{\partial\phi}{\partial t}\right]. \qquad (33)$$

Combining equations (30) and (32) the integral to be minimized has the form

$$I = \frac{\int \phi H\phi\, d\tau}{\int \phi^2 d\tau} = \frac{\int \tfrac{1}{2}\{(\mathrm{grad}_1\phi)^2 + (\mathrm{grad}_2\phi)^2\}d\tau + \int V\phi^2 d\tau}{\int \phi^2 d\tau}, \qquad (34)$$

where $d\tau = dxdydz$.

According to equation (29), V is given by the relation

$$V = -\frac{Z}{r_1} - \frac{Z}{r_2} + \frac{1}{r_{12}}$$

$$= -\frac{4Zus - s^2 + t^2}{u(s^2 - t^2)}. \tag{35}$$

From equations (33), (34), and (35) it follows that the integral to be minimized has the form

$$I = \frac{M - L}{N} = \lambda, \tag{36}$$

where λ is the eigenvalue which appears in (12) and (17).

$$M = \int_0^\infty ds \int_0^s du \int_0^u dt \left\{ u(s^2 - t^2) \left[\left(\frac{\partial \phi}{\partial s}\right)^2 + \left(\frac{\partial \phi}{\partial t}\right)^2 + \left(\frac{\partial \phi}{\partial u}\right)^2 \right] + \frac{2\partial \phi}{\partial u} \left[s(u^2 - t^2) \frac{\partial \phi}{\partial s} + t(s^2 - u^2) \frac{\partial \phi}{\partial t} \right] \right\}, \tag{37}$$

$$L = \int_0^\infty ds \int_0^s du \int_0^u dt \phi^2 (4Zus - s^2 + t^2), \tag{38}$$

and

$$N = \int_0^\infty ds \int_0^s du \int_0^u dt u(s^2 - t^2)\phi^2. \tag{39}$$

Owing to the fact that $\phi^2(s, t, u) = \phi^2(s, -t, u)$, and $\phi H \phi(s, t, u) = \phi H \phi(s, -t, u)$, the limits of integration have been taken as $0 \leqslant t \leqslant u \leqslant s \leqslant \infty$, and this leads to the elimination of a factor 2 in the expressions for M, L, and N.

Equation (36) may be written in the form

$$\delta(M - L - N\lambda) = 0, \tag{40}$$

indicating that the expression in the brackets is to be a minimum.

We now have to consider different forms of the expression for ϕ, and the simplest form is that chosen in the previous section, *viz.*:

$$\phi = \epsilon^{-k(r_1 + r_2)} = \epsilon^{-ks}, \tag{41}$$

where, for the present, the normalization factor may be omitted, and k is a parameter, the value of which has to be determined from the condition (40).

Since

$$\int_0^\infty ds \int_0^s du \int_0^u u\,dt(s^2 - t^2) = \int_0^\infty ds \int_0^s du \cdot u\left(s^2 u - \frac{u^3}{3}\right)$$

$$= \int_0^\infty ds \left(\frac{s^5}{3} - \frac{s^5}{15}\right) = \int_0^\infty ds \cdot \frac{4}{15} s^5,$$

while

$$\int_0^\infty ds \int_0^s du \int_0^u dt(4Zus - s^2 + t^2)$$

$$= \int_0^\infty ds \int_0^s du \left\{(4Zus - s^2)u + \frac{u^3}{3}\right\}$$

$$= \int_0^\infty ds \left\{\frac{4Z}{3} - \frac{5}{12}\right\} s^4,$$

and

$$\frac{\partial \phi}{\partial t} = \frac{\partial \phi}{\partial u} = 0,$$

equation (40) reduces to

$$\delta \int_0^\infty ds \left\{\frac{4s^5}{15}\left(\frac{d\phi}{ds}\right)^2 - \left(\frac{4Z}{3} - \frac{5}{12}\right) s^4 \phi^2 - \frac{4\lambda s^5}{15}\phi^2\right\} = 0. \tag{42}$$

The Euler equation which this variation problem must satisfy is

$$\frac{d}{ds} \cdot \frac{\partial F}{\partial\left(\frac{d\phi}{ds}\right)} - \frac{\partial F}{\partial \phi} = 0, \tag{43}$$

where F denotes the function in the large brackets in equation (42).

Since

$$\frac{\partial F}{\partial\left(\frac{d\phi}{ds}\right)} = \frac{8s^5}{15} \cdot \frac{d\phi}{ds}$$

and

$$\frac{\partial F}{\partial \phi} = -2s^4\left(\frac{4Z}{3} - \frac{5}{12} + \frac{4\lambda s}{15}\right)\phi,$$

therefore the Euler condition (43) corresponds to the differential equation

$$\frac{8s^5}{15}\frac{d^2\phi}{ds^2} + \frac{8s^4}{3}\frac{d\phi}{ds} + 2s^4\left(\frac{4Z}{3} - \frac{5}{12} + \frac{4\lambda s}{15}\right)\phi = 0.$$

That is,

$$\frac{d^2\phi}{ds^2}+\frac{5}{s}\frac{d\phi}{ds}+\left(\lambda+\frac{5Z}{s}-\frac{25}{16s}\right)\phi=0,$$

which is the corresponding S. equation.

Substituting for ϕ from equation (41), the result is

$$k^2-\frac{5k}{s}+\lambda+\frac{5Z}{s}-\frac{25}{16s}=0.$$

Since this relation must be valid for all values of s, we must have the following two relations:

$$k = Z-\tfrac{5}{16} = \tfrac{27}{16}, \text{ since } Z=2$$

and

$$\lambda = -k^2 = -(\tfrac{27}{16})^2.$$

Hence,

$$E = \lambda\cdot 2Rch$$

$$= -\tfrac{729}{512}RchZ^2,$$

and

$$\phi = \epsilon^{-\frac{27}{16}(r_1+r_2)}$$

$$= \epsilon^{-\frac{27}{32}Z(r_1+r_2)},$$

which is the same form as equation (14), except for the normalization factor.

As mentioned previously, the value thus derived for V_i, the ionization energy of the helium atom, is 1.53 volts less than the observed value. This difference is due to the fact that the assumed eigenfunction leads to a distribution function for each electron in which the repulsive force between the electrons is not taken into account to a sufficient extent. The electrons do not move independently, and actually there must exist a high degree of improbability for the simultaneous occurrence of the two electrons in adjacent regions in the three-dimensional space. This phenomenon gives rise to an energy term which E. Wigner and F. Seitz have designated as *correlation energy*.

In order to take this into account it is necessary to use functions which involve both t and u as variables. These functions should exhibit minima in the region $u = 0$ and t very large. Thus the next approximation used by Hylleraas has the form

$$\phi = \epsilon^{-ks}(1 + a_1u + a_2t^2),$$

where k, a_1, and a_2 are variable parameters which must satisfy the three conditions

$$\frac{\partial\lambda}{\partial k}=\frac{\partial\lambda}{\partial a_1}=\frac{\partial\lambda}{\partial a_2}=0.$$

This yields a value for V_i which is only 0.033 volt less than the observed value.

The next approximation is obtained by means of the function

$$\phi = \epsilon^{-ks}(1 + a_1u + a_2t^2 + a_3s + a_4s^2 + a_5u^2),$$

which reduces the difference to 0.0115 volt.

To obtain the most accurate approximate, Hylleraas has used the function

$$\phi = \epsilon^{-ks}P(s, t, u),$$

where

$$P = \sum_{n,l,m} A_{n,l,m}s^n t^{2l}u^m,$$

that is, where ϕ is expressed as a series of terms in powers of s, t, and u. Under these conditions it is necessary to solve a number of equations of the form

$$\frac{\partial\lambda}{\partial A_{n,l,m}} = 0.$$

By using a polynomial with fourteen terms, a value was deduced for V_i which is within 0.0016 volt of the spectroscopic.

11.4 The Problem of the Many-Electron Atom. The discussion in the previous sections indicates that, even in the case of the two-electron system, the problem of solving the S. equation involves a great deal of tedious calculation, if results of any considerable degree of accuracy are required. It is obvious that the difficulties encountered in solving the corresponding equation in the case of the many-electron atom must be even greater. In the following section some of the more outstanding methods which have been developed for the solution of this very complex problem are described quite briefly. The reader who desires to follow up this aspect of quantum mechanics will find more complete details in the references given at the end of the chapter.

In any atomic system we may divide the electrons into two classes: (1) those in "inner shells," and (2) those in the outer shell, or valence group. It is only the latter which we need to consider, in general, in calculating the energy of the system, or the mode of interaction of one atom with other atoms.

For these electrons we can apply the variational principle in the form

$$\int \bar{\phi}H\phi d\tau = E^0 \int \bar{\phi}\phi d\tau,$$

where E^0 is the zero-order eigenvalue expressed in ordinary units. The Hamiltonian operator for an atom with N valence electrons has the

form

$$H = -\sum_{i=1}^{N}\left(\frac{1}{\alpha^2}\nabla_i^2 + \frac{Ze^2}{r_i}\right) + \sum_{i,j}{}' \frac{e^2}{r_{ij}}, \tag{44}$$

where $\alpha^2 = 8\pi^2\mu/h^2$, and the prime over the last summation indicates that the combination i, j is taken only for values of $j > i$.

The number of terms in this summation is equal to the number of combinations of N things, taken two at a time, that is,

$$C_2(N) = \frac{N!}{(N-2)!2!}.$$

For the case $N = 2$, the last equation evidently becomes identical with equation (*10*.10*a*). If the " perturbation " terms in $1/r_{ij}$ were absent, the solution would be the product of N eigenfunctions, one for each electron. Hence, the important problem in the solution of equation (44) is that of taking into account the repulsive forces between the electrons. In order to realize better the conditions which must be satisfied by any distribution function which shall adequately represent the behavior of the electrons in the system, it is necessary to consider more fully the significance of the terms in the expression for H.

This may be separated into two parts: one corresponding to the kinetic energies of the particles, and the other representing the potential energy function. These are given by the relations

$$T = -\sum_i \frac{1}{\alpha^2}\nabla_i^2, \tag{45}$$

and

$$V = \sum_i V_i + \sum_{i,j}{}' \frac{e^2}{r_{ij}}, \tag{46}$$

where

$$V_i = -\frac{Ze^2}{r_i}.$$

As in section 7.8, the mean value of the kinetic energy may be written in the form

$$\bar{T} = \int \bar{\phi} T \phi d\tau$$

$$= \frac{1}{\alpha^2}\sum_i \int |\operatorname{grad}_i \phi|^2 d\tau_i, \tag{47}$$

and the mean value of V is given by

$$\bar{V} = \int |\phi|^2 V d\tau. \tag{48}$$

The latter integral may be regarded as the classical potential energy of the charge distribution $\rho = |\phi|^2$ in the potential field V in the configuration space of $3N$ dimensions.

Since the integral in (48) is negative in the case of stable states of atomic systems, the mean value of V is always negative, while that of T is positive. In order to obtain a minimum value (that is, a maximum negative value) for E, the total energy, it is therefore necessary to obtain such a representation for $|\phi|^2$ as will give a minimum value for $\bar{T}$ and maximum negative value for $\bar{V}$.

Now we can make $\bar{T}$ very small, by choosing such a form for ϕ as will make each of the functions $|\text{grad}_i\phi|^2$ quite small. This means that ϕ *changes only slowly* with variation in the coördinate variables, and therefore the distribution is spread out over a large region. The same conclusion is also evident from the point of view of the Principle of Indeterminacy. A high degree of uncertainty in the position of a particle corresponds to a fairly definite knowledge of the momentum of the particle. But such a distribution will overemphasize the effect of large values of r_i in the expression for $\bar{V}$, with the result that the latter will not have as large a negative value as it should have. Consequently E will not be as negative as possible.

On the other hand, if the electrons are localized to any extent, this must correspond to very high rates of change in ϕ with variation in the coördinates. That is, $\bar{T}$ is increased enormously, while $\bar{V}$ is made highly negative by letting $|\phi|^2$ have very large values in regions where r_i is small. This conclusion also follows from a consideration of the Principle of Indeterminacy, since precise knowledge of position can occur only if the momentum can vary over an extremely large range.

When we consider the effect of the $1/r_{ij}$ terms, it is evident from equation (46) that, if the electrons are localized in adjacent regions (for which r_{ij} is very small), the resulting value of $\bar{V}$ is made more positive, and E must therefore become more positive. Consequently, it is necessary to choose such a form for ϕ as will make the terms involving $1/r_{ij}$ as small as possible. That is, the best form of distribution function will be that which represents each electron as tending to avoid the particular region in which any other electron is present.

These remarks are equally valid in the quantum mechanics consideration of the problems of molecule formation and of the solid state. It is the necessity for satisfying apparently opposing conditions that makes the actual solution of the S. equation difficult in most of the cases which are of practical interest.

Of the methods which have been developed for treating the problem of the many-electron atom, that of the self-consistent field developed by

D. R. Hartree[12] and that involving antisymmetrical functions utilized by J. C. Slater[13] are the most important.

11.5 The Method of the Self-Consistent Field. The eigenfunction for the atomic system is represented as a product of single-electron eigenfunctions. These are obtained by a method which may be described best by considering its application in the case of the helium atom.[14]

Let us assume that the potential energy function can be expressed as the sum of two terms in the form

$$V = V_1(r_1) + V_2(r_2),$$

where V_1 is a function of r_1, θ_1, and η_1, the coördinates of one electron, and V_2 is a function of the coördinates of the second electron. From a knowledge of these functions it should be possible to determine the eigenfunction for the system as the product to two single-electron functions $\phi_1(r_1)$ and $\phi_2(r_2)$. But, in that case, the potential field effective for electron 1 is given (in Hartree units) by

$$V_1(r_1) = -\frac{Z_1}{r_1} + \int |\phi_2(r_2)|^2 \frac{d\tau_2}{r_2}, \tag{49}$$

where the first term represents the attraction due to the nucleus, and the second term, the repulsion due to the *averaged* charge distribution for electron 2.

That is, it is possible to calculate $V_1(r_1)$ by utilizing a plausible form for $\phi_2(r_2)$. But, $V_1(r_1)$ having been determined, it is possible to solve the S. equation

$$\nabla^2\phi_1(r_1) + \{E_1 - V_1(r_1)\}\phi_1(r_1) = 0 \tag{50}$$

for the function $\phi_1(r_1)$. This can then be inserted in an equation similar to (49) for the determination of $V_2(r_2)$, and gives the average potential field effective for electron 2. The function thus obtained can be inserted in a S. equation similar to (50) for the determination $\phi_2(r_2)$. The form of the function derived by this procedure should agree with that assumed in solving equation (49). From the extent to which this agreement is actually obtained by a first trial form for $\phi_2(r_2)$, it is possible to decide upon an improved form for the function, and the procedure is then repeated with the latter. This explains the

[12] D. R. Hartree, *Proc. Cambridge Phil. Soc.*, **24**, 89, 111, 426 (1928), and subsequent papers in *Proc. Roy. Soc. (London)*, A.

[13] J. C. Slater, *Phys. Rev.*, **34**, 1293 (1929).

[14] This illustration is given by N. F. Mott, "Wave Mechanics," p. 120. See also H. Bethe, "Handbuch der Physik," XXIV/1, pp. 368–371.

reason for designating Hartree's method as that of *the self-consistent field.*

V. Fock has shown[15] that if, in the general case, a product function

$$\phi = \phi_1(r_1)\phi_2(r_2) \ldots \phi_n(r_n)$$

is chosen and the functions $\phi_1(r_1)$, etc., are varied individually to obtain the minimum value of the variational integral $\int \phi H \phi d\tau$, then the same single-electron functions are derived as by the application of Hartree's method.

J. C. Slater[16] has also examined Hartree's method critically as to its accuracy as a method of solving the S. equation and has pointed out the conditions under which the eigenfunctions and energy values thus deduced might deviate to a considerable extent from the correct values. However, he has also shown[17] that in the case of the normal helium atom the method yields an electron distribution function which agrees well with that derived independently by Hartree and an ionization energy which is within 1 per cent of the observed value. Furthermore, Hartree and his associates have applied the method to calculate distribution functions for electrons in more complex atoms and have obtained results which are in satisfactory agreement with deductions from observations on the scattering of X-rays and of electrons by atoms.

11.6 Slater's Treatment of the Many-Electron Problem. In accordance with Pauli's Exclusion Principle the *complete eigenfunction* for any atomic or molecular system must be antisymmetric. Consequently Hartree's function is not quite satisfactory. However, Slater has shown that an antisymmetric function may be built up out of single-electron functions in the following manner.

Let $\phi_{n_i}(x_i, y_i, z_i)$ denote the eigenfunction for the electron having the quantum numbers n_i, l_i, m_i and the coördinates of position x_i, y_i, z_i. The spin function will be designated by $\alpha_i(\omega_i)$, where ω_i may have the values $\pm(\frac{1}{2})h/(2\pi)$. The complete eigenfunction is given by

$$u(n_i/x_i) = \phi_{n_i}(x_i, y_i, z_i) \cdot \alpha_i(\omega_i),$$

where n_i designates the four quantum numbers n, l, m, and m_s, and x_i designates the four coördinates, three of position and one of spin.

A Hartree function would be obtained by taking the product of similar

[15] *Z. Physik*, **61**, 126 (1930). See also Pauling and Wilson, "Quantum Mechanics," p. 252.

[16] *Phys. Rev.*, **32**, 339 (1928).

[17] *Phys. Rev.*, **32**, 349 (1928).

functions for each of the N electrons in the system. This would have the form

$$u(n_1/x_1)\, u(n_2/x_2) \ldots u(n_N/x_N),$$

and it satisfies the S. equation approximately. This, however, does not satisfy the exclusion principle. As Slater points out, "To build up an antisymmetric solution, we first note that we still have an approximate solution, connected with the same energy value, if we interchange any two x's, obtaining for example $u(n_1/x_2)\, u(n_2/x_1) \ldots u(n_N/x_N)$. We still have an approximation with the same energy if we make a linear combination of any such solutions. Then we can make the one possible combination which is antisymmetric, and it will both satisfy the exclusion principle, and will be an approximate solution of Schrödinger's equation."

This combination is written by Slater in the form of the determinant:

$$\phi_0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} u(n_1/x_1) & u(n_1/x_2) & \ldots & u(n_1/x_N) \\ u(n_2/x_1) & u(n_2/x_2) & \ldots & u(n_2/x_N) \\ \cdots & \cdots & \cdots & \cdots \\ u(n_N/x_1) & u(n_N/x_2) & \ldots & u(n_N/x_N) \end{vmatrix}, \tag{51}$$

where $1/\sqrt{N!}$ is the normalizing factor, if each of the individual functions is normalized.

It is obviously antisymmetric, for interchanging, say, x_1 and x_2 interchanges two columns of the determinant, which by a familiar property merely changes the sign. It can be shown that it is the only antisymmetric combination of these functions, and it leads at once to the familiar interpretation of the exclusion principle. For, if two of the functions had the same quantum numbers (say $n_1 = n_2$ symbolizing equality of the four quantum numbers), then the corresponding rows of the determinant would be identical [since they contain the functions $u(n_1/x_1) = u(n_2/x_1) = u(n_2/x_2)$, etc.], and by another familiar rule the determinant would vanish. Thus, there is no solution corresponding to the case where two electrons have the same set of quantum numbers. Further, the determinant treats all electrons alike; hence we cannot count as separate two states which differ only by the interchange of the quantum numbers of two electrons. Our exclusion principle then coincides with the one previously described.

The Slater determinant can also be written in the form

$$\phi_0 = \frac{1}{\sqrt{N!}} \sum_P (-1)^P P u(n_1/x_1)\, u(n_2/x_2) \ldots u(n_N/x_N), \tag{52}$$

in which P denotes any permutation of the electron coördinates, and

$(-1)^P = +1$ when P is even, that is, is equivalent to an even number of interchanges, while $(-1)^P = -1$, when P is *odd*.

As Pauling and Wilson observe,[18] "The function ϕ_0 takes care of the degeneracy due to the $N!$ possible distributions of the N electrons in a fixed set of N functions, u. There still remains another type of degeneracy, due to the possibility of there being more than one set of spin-orbit functions corresponding to the same unperturbed energy." For these cases, a Slater determinant has to be set up for each set of spin coördinates, and the complete eigenfunction will be represented by the sum of two or more Slater determinants. The energy values are then obtained by solving a secular equation.

Returning to the case in which only one Slater determinant is required, the energy is given by the relation

$$E^0 = \int \phi_0 H \phi_0 d\tau d\omega, \tag{53}$$

where $d\tau$ is the element of configuration space and $d\omega$ that of spin space. Assuming no interaction between spin and orbital functions, the last integral can be written as the product of two integrals, one involving $d\tau$, the other $d\omega$. Let $\alpha_i(\omega_i)$ denote the function for one direction of spin, and $\beta_i(\omega_i)$ that for the opposite direction. Then we have the relations

$$\left.\begin{aligned} \int \alpha_i^2(\omega_i) d\omega_i &= \int \beta_i^2(\omega_i) d\omega_i = 1 \\ \int \alpha_i(\omega_i)\beta_i(\omega_i) d\omega_i &= 0. \end{aligned}\right\} \tag{54}$$

In the case of the integral in (53) it is evident, since H does not operate on the spin functions, that the spin functions will occur as α_i^2 or β_i^2. Consequently, the integral involving $d\omega$ is equal to unity, and we need to consider only the integral over the configuration space.

Now let us consider the integral $\int \phi_0 H \phi_0 d\tau$. This may also be written in the form

$$E^0 = -\frac{1}{\alpha^2} \int \phi_0 \nabla^2 \phi_0 d\tau + \int V \mid \phi_0 \mid^2 d\tau. \tag{55}$$

Since each eigenfunction is the solution of a S. equation of the form

$$\frac{1}{\alpha^2} \nabla^2 u_i + (E_i - V_i) u_i = 0,$$

[18] "Introduction to Quantum Mechanics," p. 233.

the first integral in (55) can be reduced to the sum of a series of integrals each of which is equal to $E_i - \int u_i^2 V_i d\tau_i$.

The second integral in (55) will consist of two types of terms of which the first has the form

$$J_{12} = \int \frac{\{u(n_1/x_1)\}^2 \{u(n_2/x_2)\}^2}{r_{12}} d\tau_1 d\tau_2, \tag{56}$$

and the second has the form

$$K_{12} = \int \frac{u(n_1/x_1)\, u(n_2/x_2)\, u(n_1/x_2)\, u(n_2/x_1)}{r_{12}} d\tau_1 d\tau_2. \tag{57}$$

The integral J_{12} represents the *Coulomb interaction*, while K_{12} corresponds to *exchange interaction*. It is evident that these exchange integrals are *not due to any new type of interaction between the electrons*, but arise from the fact that ϕ_0 is expressed in the form of a determinant.

As mentioned already, the degenerate case requires, for the complete solution, a number of determinantal functions, each corresponding to a definite set of spin functions. Let ϕ_α, ϕ_β, . . . designate these functions. It is then necessary to solve a secular equation of which the matrix elements have the form

$$\phi_\alpha H \phi_\beta - \delta_{\alpha\beta} E,$$

where $\delta_{\alpha\beta} = 1$ or 0.

The actual computation of the energy levels in any given case is quite tedious, and, in view of the fact that complete details and illustrations of such calculations are given both by Slater and by Pauling and Wilson, it has not been considered necessary to discuss this topic at further length in the present connection. Moreover, there will be given in a subsequent chapter a calculation, based on the same methods, of the energy of interaction of two or more atoms, which will also serve to illustrate the principles involved.

In concluding this brief summary of the methods which have been used for solving the problems of many-electron atoms mention should be made of the modification of Slater's method introduced by V. Fock,[19] and of the publications by J. E. Lennard-Jones[20] in which the methods of the perturbation theory are applied.

[19] *Z. Physik*, **61**, 126 (1930).

[20] *Proc. Roy. Soc.* (*London*), **A129**, 598 (1930), and *Proc. Cambridge Phil. Soc.*, **27**, 469 (1931).

In a more recent paper by P. M. Morse, L. A. Young, and E. S. Haurwitz[21] the variational method has been applied to calculate energy levels for a number of the simple atoms (He, Li, Be, B, C, N, O, F, Ne, Na, and Mg). The forms of single-electron functions used are as follows:

$$1s \ldots u_1(r) = \left(\frac{\mu^3 a^2}{\pi}\right)^{\frac{1}{2}} \epsilon^{-\mu a r}$$

$$2s \ldots u_2(r) = \left(\frac{\mu^5}{3\pi N}\right)^{\frac{1}{2}} \left\{ r\epsilon^{-\mu r} - \frac{3A}{\mu} \epsilon^{-\mu b r} \right\}$$

$$2p \ldots u_3(r) = \left(\frac{\mu^5 c^5}{\pi}\right)^{\frac{1}{2}} r \cos\theta \, \epsilon^{-\mu c r}$$

$$u_4(r) = \left(\frac{\mu^5 c^5}{2\pi}\right)^{\frac{1}{2}} r \sin\theta \, \epsilon^{i\eta - \mu c r}$$

$$u_5(r) = \left(\frac{\mu^5 c^5}{2\pi}\right)^{\frac{1}{2}} r \sin\theta \, \epsilon^{-i\eta - \mu c r}.$$

The constant A is fixed so that u_2 is orthogonal to u_1. Its value is given in terms of the two parameters a and b, and N is a normalizing factor which is also a function of a and b. "The parameter μ is a scale factor, whose best value can be determined analytically, leaving but three parameters to be determined numerically." These parameters a, b, and c are then determined from the conditions

$$\frac{\partial E}{\partial a} = \frac{\partial E}{\partial b} = \frac{\partial E}{\partial c} = 0,$$

which lead to minimum values for E.

Tables are given in the original paper for the calculation of the corresponding Coulomb and exchange terms, thus facilitating the calculation for any given electron configuration, when spin degeneracy is not taken into account. For calculating multiplet levels it is therefore necessary to form linear combinations of Slater determinants as mentioned in a previous section.

[21] *Phys. Rev.*, **48**, 948 (1935).

SUPPLEMENTARY NOTE 1

THE EXPRESSION FOR THE OPERATOR *H* IN ATOMIC UNITS

For an atomic system of nuclear charge Z and N valence electrons, the S. equation in terms of atomic units has the form, which is analogous to equation (12),

$$\sum_i^N \nabla^2\phi + 2\left(\lambda + \sum_i^N \frac{Z}{r_i} - \sum_{i,j}^{\prime} \frac{1}{r_{ij}}\right)\phi = 0,$$

where the summation symbols have the same significance as in equation (44).

Since the eigenvalue λ for this system must satisfy the condition

$$\int \bar{\phi} H \phi d\tau = \lambda \int \bar{\phi}\phi d\tau,$$

it follows that the Hamiltonian operator is given, in terms of atomic units, by the relation

$$H = -\frac{1}{2}\sum_i^N \nabla^2 - \sum_i^N \frac{Z}{r_i} + \sum_{i,j}^{\prime} \frac{1}{r_{ij}}.$$

Thus the introduction of the factor $\frac{1}{2}$ in equation (30) is due to the particular choice of units in which the energy is expressed.

COLLATERAL READING

1. For discussion of the variational method see the following:

   BETHE, H., " Handbuch der Physik," XXIV/1, p. 354.
   PAULING, L., and WILSON, E. B., " Introduction to Quantum Mechanics," Chapter VII.

2. The problem of the many-electron atom:

   PAULING, L., and WILSON, E. B., Chapter IX.
   SLATER, J. C., *Phys. Rev.*, **34**, 1293 (1929).

CHAPTER XII

THE HYDROGEN MOLECULE

12.1 Heitler-London Method. The problem of determining the solution of the S. equation for the helium atom is a limiting case of the problem which is presented by the hydrogen molecule. If in the latter we permit the two nuclei to approach until they coalesce, we obviously obtain the same system as that of the helium atom. Consequently, we expect to find in the solution of the S. equation for the hydrogen molecule certain similarities with the results obtained in the previous chapter. However, there is one feature of the hydrogen molecule which is of extreme importance, and which does not occur in the case of the helium atom. This concerns the interpretation of the valence bond, which, on the basis of the Lewis-Langmuir theory, is regarded as due to the sharing of two electrons between the two hydrogen nuclei. The first successful application of quantum mechanics to the solution of the problem of the hydrogen molecule was made by W. Heitler and F. London,[1] and although subsequent investigators have developed methods of attacking the problem by which more accurate results have been obtained, we shall find it advantageous to discuss the Heitler-London (HL) theory in some detail before describing some of the other methods.

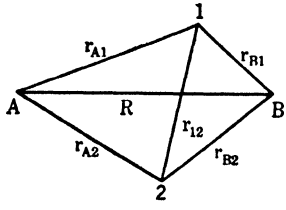


FIG. 54. Illustrating notation used in formulating potential energy function for interaction of two hydrogen atoms.

We consider a system consisting of two nuclei A and B, and two electrons 1 and 2, as indicated in Fig. 54. In the unperturbed state, where the two atoms are quite separate, it is obvious that the energy $E = 2E_0$, where E_0 is the energy of the hydrogen atom in the normal state, as given in Chapter VII. For electron 1 attached to nucleus A and electron 2 to nucleus B, the eigenfunction is given by

$$\phi_1 = u_A(1)u_B(2), \tag{1}$$

[1] Heitler and London, *Z. Physik*, **44**, 455 (1927).

where

$$\left.\begin{aligned} u_A(1) &= -\frac{1}{\sqrt{\pi a_0^3}} \cdot \epsilon^{-\frac{r_{A1}}{a_0}} \\ u_B(2) &= -\frac{1}{\sqrt{\pi a_0^3}} \cdot \epsilon^{-\frac{r_{B2}}{a_0}} \end{aligned}\right\}, \tag{2}$$

and r_{A1} and r_{B2} refer to distances indicated in Fig. 54.

But corresponding to the same energy $2E_0$, it is also possible to have the eigenfunction obtained by interchanging the electrons, that is,

$$\phi_2 = u_A(2)u_B(1) \tag{3}$$

$$= \frac{1}{\pi a_0^3} \cdot \epsilon^{-\frac{(r_{A2}+r_{B1})}{a_0}}. \tag{4}$$

Consequently, the system is *degenerate* and the zero-order eigenfunction should be represented, as in the case of the excited helium atom, by the two linear combinations

$$\phi_\alpha^0 = a\phi_1 + b\phi_2, \tag{5}$$

$$\phi_\beta^0 = c\phi_1 + d\phi_2, \tag{6}$$

where a, b, c, and d are constants.

The eigenfunctions representing to first-order perturbation terms, the perturbed state, that is, the state in which the two atoms are interacting, will be given by

$$\phi_\alpha = \phi_\alpha^0 + \psi_\alpha, \tag{7}$$

and

$$\phi_\beta = \phi_\beta^0 + \psi_\beta, \tag{8}$$

while the corresponding eigenvalues will be

$$E_\alpha = 2E_0 + \eta_\alpha, \tag{9}$$

and

$$E_\beta = 2E_0 + \eta_\beta. \tag{10}$$

These will represent the eigenfunctions and eigenvalues corresponding to the S. equation:

$$\nabla_1^2\phi + \nabla_2^2\phi + \kappa\,[E - V_0 + V_1^A + V_2^B + V_1^B + V_2^A]\,\phi = 0, \tag{11a}$$

or

$$H\phi = E\phi, \tag{11b}$$

where the *operator* H is defined by the relation

$$H \equiv \left\{ -\frac{1}{\kappa}(\nabla_1^2 + \nabla_2^2) + V_0 - (V_1^{\mathrm{A}} + V_2^{\mathrm{B}} + V_1^{\mathrm{B}} + V_2^{\mathrm{A}}) \right\}, \tag{12}$$

and

$$\left.\begin{aligned} \kappa &= \frac{8\pi^2\mu}{h^2} \\ V_0 &= \frac{e^2}{R} + \frac{e^2}{r_{12}} \\ V_1^{\mathrm{A}} = \frac{e^2}{r_{\mathrm{A}1}},\ V_2^{\mathrm{B}} &= \frac{e^2}{r_{\mathrm{B}2}},\ V_1^{\mathrm{B}} = \frac{e^2}{r_{\mathrm{B}1}},\ V_2^{\mathrm{A}} = \frac{e^2}{r_{\mathrm{A}2}} \end{aligned}\right\}. \tag{13}$$

The suffixes on the Laplacian operators refer to electrons 1 and 2. Since ϕ_α^0 and ϕ_β^0 must be orthonormalized functions, it follows that

$$\int (\phi_\alpha^0)^2 dv_1 dv_2 = \int (\phi_\beta^0)^2 dv_1 dv_2 = 1, \tag{14}$$

and

$$\int \phi_\alpha^0 \phi_\beta^0 dv_1 dv_2 = 0, \tag{15}$$

where the integration is carried out over the whole configuration space for each electron.

By substituting in these last two equations from equations (5) and (6), it follows that

$$a^2 + b^2 + 2abS^2 = 1, \tag{16}$$

$$c^2 + d^2 + 2cdS^2 = 1, \tag{17}$$

$$ac + bd + (ad + bc)S^2 = 0, \tag{18}$$

where

$$\begin{aligned} S^2 &= \int \phi_1 \phi_2 dv_1 dv_2 \\ &= \int u_{\mathrm{A}}(1) u_{\mathrm{B}}(2) u_{\mathrm{A}}(2) u_{\mathrm{B}}(1) dv_1 dv_2. \end{aligned} \tag{19}$$

Now it is evident that for $R = \infty$, that is, for the unperturbed state, $S = 0$, since *either* ϕ_1 *or* ϕ_2 represents the two separate atoms. On the other hand, for $R = 0$, as mentioned already, the system becomes identical with that of the normal helium atom, for which

$$S^2 = \int \left\{ u(1)u(2) \right\}^2 dv_1 dv_2 = 1.$$

Hence, in the case of the hydrogen molecule, $1 > S > 0$, and since in the case of helium, the zero-order eigenfunctions for the excited states are

$$\phi_S^0 = \frac{1}{\sqrt{2}}\Big\{u_n(1)u_m(2) + u_n(2)u_m(1)\Big\},$$

and

$$\phi_A^0 = \frac{1}{\sqrt{2}}\Big\{u_n(1)u_m(2) - u_n(2)u_m(1)\Big\},$$

we assume $a = b$. The validity of this assumption will be justified by the resulting deductions.

From equation (16) it follows that

$$b = a = \frac{1}{\sqrt{2(1 + S^2)}}. \tag{20}$$

Also, from equation (18), it follows that

$$b(c + d)\ (1 + S^2) = 0.$$

Since

$$b \neq 0 \quad \text{and} \quad S^2 \neq -1, c = -d.$$

Hence, by substituting in equation (17), we deduce the result

$$c = \frac{1}{\sqrt{2(1 - S^2)}}; \quad d = \frac{-1}{\sqrt{2(1 - S^2)}}; \tag{21}$$

so that the orthonormalized functions are as follows:

$$\phi_\alpha^0 = \frac{1}{\sqrt{2(1 + S^2)}}(\phi_1 + \phi_2); \tag{22}$$

$$\phi_\beta^0 = \frac{1}{\sqrt{2(1 - S^2)}}(\phi_1 - \phi_2). \tag{23}$$

It is seen that ϕ_α^0 is a *symmetrical*, and ϕ_β^0 an *antisymmetrical, function.*

By direct substitution of these relations in equations (14) and (15) it is readily shown that the latter conditions are satisfied, so that equations (22) and (23) represent orthogonal and normalized functions.

From symmetry considerations it is also evident that

$$\int u_A(1)u_B(1)dv_1 = \int u_A(2)u_B(2)dv_2 = S. \tag{24}$$

Furthermore, these four eigenfunctions satisfy the four S. equations:

$$\left.\begin{aligned}\nabla^2 u_A(1) + \kappa(E_0 + V_1^A)u_A(1) &= 0\\ \nabla^2 u_B(2) + \kappa(E_0 + V_2^B)u_B(2) &= 0\\ \nabla^2 u_B(1) + \kappa(E_0 + V_1^B)u_B(1) &= 0\end{aligned}\right\}, \tag{25}$$

and

$$\nabla^2 u_A(2) + \kappa(E_0 + V_2^A)u_A(2) = 0$$

Hence,

$$\begin{aligned}\nabla^2\phi_1 &= u_A(1)\nabla^2 u_B(2) + u_B(2)\nabla^2 u_A(1)\\ &= -\kappa(2E_0 + V_1^A + V_2^B)\phi_1,\end{aligned} \tag{26}$$

$$\nabla^2\phi_2 = -\kappa(2E_0 + V_2^A + V_1^B)\phi_2, \tag{27}$$

and

$$\begin{aligned}\nabla^2(\phi_1 + \phi_2) + \kappa\{&2E_0(\phi_1 + \phi_2) + (V_1^A + V_2^B)\phi_1\\ &+ (V_2^A + V_1^B)\phi_2\} = 0.\end{aligned} \tag{28}$$

Now if ϕ_α is a solution of the S. equation (11a), then, by substituting from equations (7) and (22), we obtain the equation

$$\begin{aligned}&\nabla^2(\phi_1 + \phi_2) + \kappa(2E_0 + \eta_\alpha - V_0 + V_1^A + V_2^B + V_1^B + V_2^A)\\ &(\phi_1 + \phi_2) + \sqrt{2 + 2S^2}\,\{\nabla^2\psi_\alpha + \kappa(2E_0 + \eta_\alpha - V_0 + V_1^A\\ &+ V_2^B + V_1^B + V_2^A)\psi_\alpha\} = 0.\end{aligned} \tag{29}$$

From (29) and (28) it follows that

$$\begin{aligned}\sqrt{2 + 2S^2}\,\{\nabla^2\psi_\alpha + \kappa(2E_0 + \eta_\alpha - V_0 + V_1^A + V_2^B + V_1^B +\\ V_2^A)\psi_\alpha\} = \kappa\{&(V_0 - \eta_\alpha)(\phi_1 + \phi_2) - (V_1^A + V_2^B)\phi_2\\ &- (V_1^B + V_2^A)\phi_1\}.\end{aligned} \tag{30}$$

This is an *inhomogeneous* partial differential equation of the same type as that encountered in equation (9.11) in connection with the calculation of the first-order perturbation energy term. In order that it may have a solution, it is necessary that the right-hand side of equation (30) should be orthogonal with respect to the solution $\phi_1 + \phi_2$ of the corresponding homogeneous differential equation.

Hence,

$$\begin{aligned}(V_0 - \eta_\alpha)\int(\phi_1 + \phi_2)^2 dv_1 dv_2 &= (V_0 - \eta_\alpha)(2 + 2S^2)\\ &= \int(V_1^A + V_2^B)\phi_2(\phi_1 + \phi_2)dv_1 dv_2 +\\ &\int(V_2^A + V_1^B)\phi_1(\phi_1 + \phi_2)dv_1 dv_2.\end{aligned} \tag{31}$$

That is,

$$\eta_\alpha = \frac{E_{11} + E_{12}}{1 + S^2} = E_{11} - \frac{E_{11}S^2 - E_{12}}{1 + S^2}, \tag{32}$$

where

$$2E_{11} = \int \{V_0(\phi_1^2 + \phi_2^2) - (V_1^A + V_2^B)\phi_2^2 - (V_2^A + V_1^B)\phi_1^2\} dv_1 dv_2, \tag{33}$$

$$2E_{12} = \int (2V_0 - V_1^A - V_2^B - V_2^A - V_1^B)\phi_1\phi_2 dv_1 dv_2, \tag{34}$$

and S^2 is defined by equations (19) and (24).

In a similar manner it may be deduced that the first-order perturbation energy term corresponding to ϕ_β is given by the relation[2]

$$\eta_\beta = \frac{E_{11} - E_{12}}{1 - S^2} = E_{11} + \frac{E_{11}S - E_{12}}{1 - S^2}. \tag{35}$$

12.2 Physical Interpretation. The terms η_α and η_β represent, to a first-order approximation, *two different values for the interaction energy of two hydrogen atoms*, and an inspection of equations (22) and (23) shows that η_α corresponds to the *symmetrical* zero-order eigenfunction ϕ_α^0, while η_β corresponds to the *antisymmetrical* function ϕ_β^0. Thus the quantum mechanics treatment of the problem leads to the conclusion that two hydrogen atoms can interact in *two different modes*. The fundamental reason for this is the fact that it is possible for the two electrons to interchange places; or, stated in more technical language, the existence of two modes of vibrations and corresponding eigenvalues is due to the degeneracy of the system in the unperturbed state. This, in turn, occurs because the two electrons are absolutely equivalent, so that it is impossible to distinguish between them. In other words, the whole argument is a logical deduction from the Principle of Indeterminacy in the sense that, when two hydrogen atoms approach each other very closely, there is no

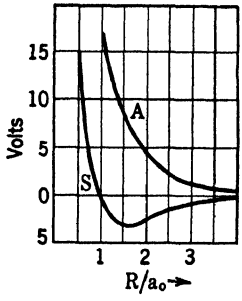


FIG. 55. Energy curves for the two modes of interaction of two hydrogen atoms; curve S corresponds to the symmetric, and curve A to the antisymmetric mode.

[2] The notation E_{11} and E_{12} is that used by Heitler and London. In a subsequent section these will be shown to correspond to matrix elements which occur in a second-degree secular equation.

experimental method by which each electron may be "tagged" and observed.

Evidently the interaction energy terms η_α and η_β are functions of R. While the actual details of evaluating the expressions for E_{11} and E_{12} are discussed in a subsequent section, Fig. 55 shows the results deduced by Y. Sugiura,[3] which are more accurate than those obtained by Heitler and London in their investigation. The values of η_α (curve S) and η_β (curve A) are given in electron-volts (v.e.)[4] as functions of the internuclear distance R/a_0, where a_0 denotes the Bohr unit radius.

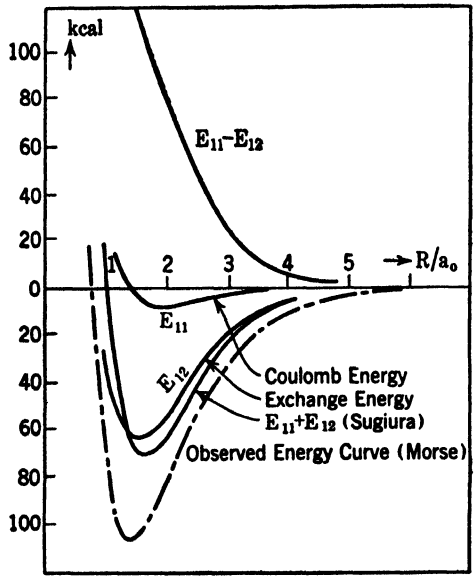


FIG. 56. Plots of the total energy, Coulomb energy and exchange energy as functions of internuclear distance, for the two modes of interaction of two hydrogen atoms.

It will be observed that η_α reaches a minimum value of -3.2 v.e. for $R = 1.52a_0 = 0.80$ Å, while η_β is always positive. Thus, η_α represents the *lower energy state* and *must correspond to molecule formation*, while η_β, since it is positive, must correspond to repulsion between the atoms.[5] Consequently, the symmetric eigenfunction ϕ_α represents a *stable state*, while the antisymmetric eigenfunction ϕ_β represents an unstable state.

Figure 56 shows plots of the same quantities in terms of calories per mole H_2. For comparison there are also plotted the energy term E_{11} and the curve calculated by P. M. Morse[6] from observations on the band spectrum of H_2. These observations lead to a minimum of the curve at $R = 1.40a_0 = 0.74$ Å. The considerations upon which this calculation has been made will be discussed in the following chapter.

Calculation shows that the term E_{12} is *negative* (over a large range of

[3] Sugiura, *Z. Physik*, **45**, 484 (1927).

[4] 1 v.e. = 23,055 cal./mole.

[5] Thus, curve S is similar to the plots shown in Fig. 51, Chapter VIII, for the energy of interaction of two molecules.

[6] Morse, *Phys. Rev.*, **34**, 57 (1929); Condon and Morse, "Quantum Mechanics," p. 163; also see the following chapter for further discussion.

values of R/a_0) and greater in absolute value than E_{11}, while S^2 is positive and less than 1 for all values of $R/a_0 > 1$. From the curves in Fig. 56 it is seen that for values of $R/a_0 \geqslant 1.5$, approximately, the relations between η_α, η_β, E_{11}, and E_{12} are those indicated diagrammatically in Fig. 57.

A consideration of equations (33) and (34) throws considerable light on the nature of the attractive energy forces which lead to molecule formation. In equation (33), the term $V_0(\phi_1^2 + \phi_2^2)$ represents physically the total repulsive energy due to the nuclear charges and the electronic charge distributions. Similarly, the other terms on the right-hand side of (33) represent the attractive energy between the nuclei and the electronic charge distributions. Hence, we may designate E_{11} as the *Coulomb interaction energy*. On the basis of classical considerations this should constitute the whole of the interaction energy between two hydrogen atoms.

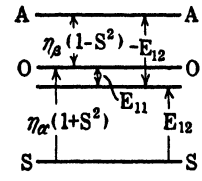


FIG. 57. Relation between the terms in the expressions for the energy of interaction of two hydrogen atoms.

Actual evaluation of the expression for E_{11} shows that this has a minimum value of -0.488 v.e. for $R = 1.90a_0 = 1.00$ Å. Since the value of S^2 for this value of R is 0.347, it follows that $-E_{11}/(1 + S^2) = 0.362$ v.e. = 8350 cal./mole. Comparing this result with the observed energy of formation of H_2, which is 4.72 v.e. = 108,900 cal./mole, it is seen that the classical electrostatic attraction and repulsion energies are quite inadequate to account for the energy of the valence bond.

As mentioned already, the energy of formation as calculated by Sugiura is 3.2 v.e. Though this is less than the observed value, nevertheless it indicates that the valence energy is accounted for to a considerable extent by the term E_{12}. What is the physical significance of this energy term? A similar expression was encountered in the exchange-energy term V_{12} which was derived in the solution of the helium atom problem. In the present case we also designate E_{12} as the *exchange energy*, and it is seen that the presence of this term is due to the assumption that the eigenfunction for the system has the form given in equations (5) or (6). In other words, E_{12} occurs because of the possibility of interchange of the electrons, and we may regard $\nu_{12} = E_{12}/h$ as a measure of the frequency of this interchange.

A great deal has been written about the non-classical nature of the term E_{12}, and since this term accounts, as shown above, for a large part of the energy of formation of H_2, a distinction has been drawn between the types of forces involved in the two energy terms (E_{11} and E_{12}). Evidently, such a distinction is only the result of the mathematical computation, for, as a matter of fact, the quantum-mechani-

cal treatment recognizes that the only forces involved in the binding of two hydrogen atoms are those which arise from electrostatic attraction and repulsion between the four particles which constitute the system. The exchange term is merely an expression of the physical requirement that the electrons in H_2 cannot be regarded as localized about the nuclei with which they were associated in the separated atoms.[7]

There is this important difference between the energy terms V_{12} and E_{12}, that, whereas E_{12} is negative in the case of the hydrogen molecule (and in most other cases of interaction of similar atoms), the term V_{12} is *positive*, as is readily evident from equation (*10*.29) and the

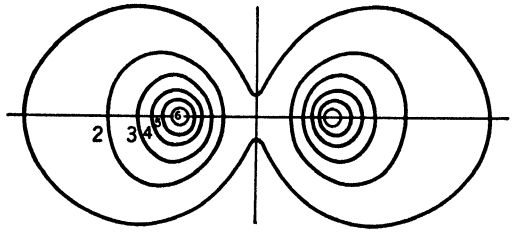


FIG. 58. Electron distribution for elastic reflection of two hydrogen atoms.

subsequent calculations in Chapter X. This is due to the fact that V_{12} represents a repulsive energy term which arises from the interchange of coördinates by two electrons of different quantum numbers in the *same atom*. On the other hand, E_{12} represents an *attractive* energy which occurs because of the possibility of interchange of two equivalent electrons, one from each of the atoms.

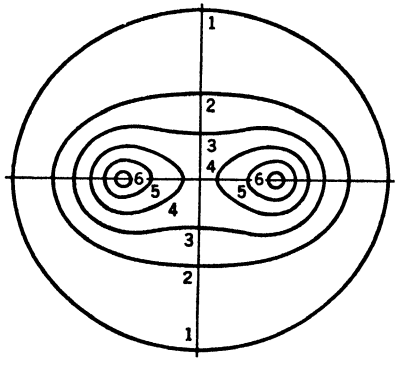


FIG. 59. Electron distribution for hydrogen molecule formation.

On the basis of these considerations Heitler and London concluded that, to a first approximation, the energy E_{12} corresponds to the energy of the *valence bond* in the molecule H_2. This is most readily evident from the plots, calculated from $(\phi_\alpha^0)^2$ and $(\phi_\beta^0)^2$, of the *density of charge distribution* for the symmetric and antisymmetric cases, respectively. Figures 58 and 59, taken from a paper by F. London,[8] show the results obtained.

The density is constant on each curve (so that they correspond to the *isobars* in atmospheric pressure measurements). The numbers

[7] S. Dushman and F. Seitz, *J. Phys. Chem.*, **41**, 233 (1937).

[8] F. London, *Leipziger Vorträge*, **1928**, pp. 59–84.

attached to each curve give the relative densities or probabilities of occurrence of the electrons. Figure 58 illustrates elastic reflection and shows that the electrons tend to occur during most of the time in the regions removed from the center. They avoid the region between the nuclei. On the other hand, Fig. 59, which represents the distribution for molecule formation (*homopolar combination*), shows that in this case the electrons tend to occur during most of the time in the *region between the nuclei.* Thus, the methods of quantum mechanics lead to an interpretation of the *shared electron pair* or *non-polar bond* of the Lewis-Langmuir theory of valency. When the chemists write the electronic formula for H_2 as H:H, they are therefore in logical agreement with the conclusion deduced from the solution of the S. equation. In other words, the theory of Heitler and London gives a quantitative basis for a representation which the chemist had derived by intuition.

12.3 The Application of the Pauli Exclusion Principle. Here, as in the case of the helium atom, we must take the electron spins into consideration. Here also we find that it is possible to obtain only four functions which are completely antisymmetrical in the electrons and which, therefore, satisfy the Pauli Exclusion Principle.

Let us designate the two spin functions $\psi(\frac{1}{2})$ and $\psi(-\frac{1}{2})$ by α and β, respectively. Then the only completely antisymmetrical functions which can be formed from the *spatial* functions ϕ_α^0 (symmetric) and ϕ_β^0 (antisymmetric) are the following:

$$\phi_\alpha^0 \frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) - \alpha(2)\beta(1)\}; \quad (\sum m_s = 0)$$

$$\phi_\beta^0 \frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) + \alpha(2)\beta(1)\}; \quad (\sum m_s = 0)$$

$$\phi_\beta^0 \alpha(1)\alpha(2) \quad \text{and} \quad \phi_\beta \beta(1)\beta(2). \quad (\sum m_s = \pm 1)$$

Only the first of these four functions involves ϕ_α^0, and in that case, $\sum m_s = 0$; that is, the spins are antiparallel. In the other three functions the total spin has the values $+1$, 0, and -1.

Thus we find that there exist three possible states in which the atoms repel each other and one state in which they attract and form a molecule. That is, when two hydrogen atoms collide there is a 25 per cent probability that this collision will result in the formation of a molecule. Furthermore, in the molecule, the spins of the two electrons must be *antiparallel.* For this reason the normal state of the molecule is designated spectroscopically as a singlet state ($^1\Sigma_g$), whereas the repulsive state is of the triplet type ($^3\Sigma_u$).

The suffixes g and u signify respectively "gerade," that is, even, and "ungerade," that is, odd. These designations refer to the fact that

$$\phi_\alpha(x_1y_1z) = +\phi_\alpha(-x, -y, -z),$$

while

$$\phi_\beta(x_1y_1z) = -\phi_\beta(-x, -y, -z).$$

Although the energy state corresponding to ϕ_β is unstable, "it is used to explain a certain continuous spectrum emitted by H_2, due to transitions from an excited triplet state, not shown in Fig. 55, down to the ${}^3\Sigma_u$ curve. Transitions down to this curve would obviously lead to a dissociation of the molecule, giving rise to a continuous spectrum."[9]

12.4 Calculation of Perturbation Energy Terms. Let us consider first the integral

$$S = \int u_A(1)u_B(1)dv_1 \tag{24}$$

$$= \frac{1}{\pi a_0^3}\int \epsilon^{-\frac{(r_{A1}+r_{B1})}{a_0}}\, dv_1. \tag{36a}$$

We shall find it advantageous to express all distances in atomic units, and designate these by ρ, with a corresponding subscript. Thus $\rho_{A1} = r_{A1}/a_0$, and so forth. In terms of these units,

$$S = \frac{1}{\pi}\int \epsilon^{-(\rho_{A1}+\rho_{B1})}dv_1. \tag{36b}$$

Now, in evaluating integrals such as those for S, E_{11}, E_{12}, and others which occur in *two-center* problems, it is found convenient to utilize *spheroidal coördinates* of the particular type known as *confocal elliptic*. These are defined as follows:

$$\lambda = \frac{1}{D}(\rho_{A1} + \rho_{B1}), \tag{37a}$$

and

$$\mu = \frac{1}{D}(\rho_{A1} - \rho_{B1}), \tag{37b}$$

where $D = R/a_0$.[10] As third coördinate, we use the angle θ, which a plane passing through the two nuclei and the instantaneous position of

[9] J. H. Van Vleck and A. Sherman, "The Quantum Theory of Valence," *Rev. Modern Phys.*, **7**, 167 (1935). This is also discussed in the following chapter.

[10] The significance of λ and μ in these equations is, of course, not to be confused with the interpretations used in other connections.

the electron makes with a fixed plane through the two nuclei. It will be recognized that λ defines an ellipse described about the two nuclei as focal points and passing through the point designating the position of electron 1. Similarly μ defines a hyperbola which passes through the same point. By revolving such a system of confocal ellipses and hyperbolas about the line joining the nuclei as axis, there is obtained a set of confocal ellipsoids and hyperboloids, which (see Fig. 61 in supplementary note 1) intersect along circles. These circles lie in planes perpendicular to the axis and have their centers on this axis. Hence, in order to specify the position of the electron, we require not only λ and μ (which define a particular circle) but also θ, the angle which a line passing through the center and the point makes with a fixed axis.

As shown in supplementary note 1, the element of volume is given by

$$d\tau = \frac{D^3}{8}(\lambda^2 - \mu^2)d\lambda d\mu d\theta. \tag{38}$$

Hence, in terms of these coördinate variables,

$$S = \frac{D^3}{8\pi}\int \epsilon^{-D\lambda}(\lambda^2 - \mu^2)d\lambda d\mu d\theta$$

$$= \frac{D^3}{4}\int \epsilon^{-D\lambda}(\lambda^2 - \mu^2)d\lambda d\mu,$$

since the limits of θ are 0 and 2π, and we can therefore integrate directly with respect to this variable.

The limits for the other variables are defined by $1 < \lambda < \infty$; $-1 < \mu < 1$. Expressing the integral as the product of one integral with respect to λ and of the other with respect to μ,

$$S = \frac{D^3}{4}\int_1^\infty \epsilon^{-D\lambda}d\lambda\left\{\int_{-1}^1 \lambda^2 d\mu - \int_{-1}^1 \mu^2 d\mu\right\}$$

$$= \frac{D^3}{4}\int_1^\infty \epsilon^{-D\lambda}\left(2\lambda^2 - \frac{2}{3}\right)d\lambda$$

$$= \frac{1}{2}\int_D^\infty \epsilon^{-x}x^2 dx - \frac{D^2}{6}\int_D^\infty \epsilon^{-x}dx,$$

where $x = D\lambda$. The expressions for these integrals are given in Appendix III. Consequently, we obtain the result

$$S = \epsilon^{-D}\{1 + D + \tfrac{1}{3}D^2\}. \tag{39}$$

Figure 60 shows a plot of S^2 as a function of D. (In the figure this ratio is designated ρ.)

We shall now consider the relation for E_{11} as given in equation (33). The first term on the right-hand side is

$$I = \int\int V_0(\phi_1^2 + \phi_2^2)dv_1dv_2 = e^2\int\int\left(\frac{1}{R} + \frac{1}{r_{12}}\right)(\phi_1^2 + \phi_2^2)dv_1dv_2,$$

and since the two electrons are equivalent it follows that we can write

$$I = \frac{2e^2}{R} + 2J_1,$$

where

$$J_1 = e^2\int\{u_A(1)u_B(2)\}^2\frac{1}{r_{12}}\,dv_1dv_2.$$

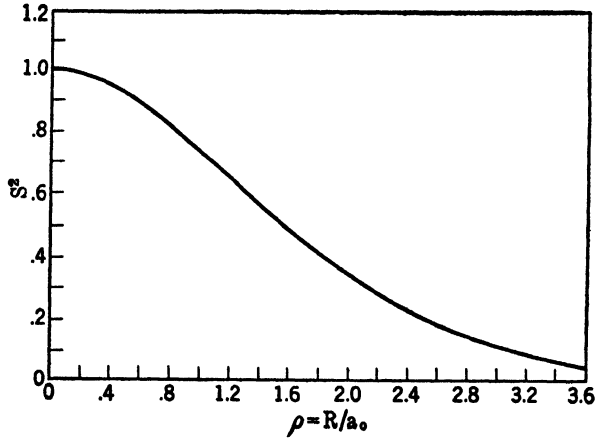


FIG. 60. Plot of the function S^2 as a function of internuclear distance.

This integral represents the *repulsive* energy between the two electronic charge distributions represented by $u_A^2(1)$ and $u_B^2(2)$, that is, by the two functions

$$u_A^2(1) = \frac{1}{\pi}\int e^{-2\rho_{A1}}\,dv_1,$$

and

$$u_B^2(2) = \frac{1}{\pi}\int e^{-2\rho_{B2}}\,dv_2.$$

Now according to equation (*10.*18), the potential at the point $\rho = \rho_{A1} = \rho_A$, due to the electron distribution *about the point* B, is given by

$$V(\rho_B) = \frac{e}{a_0\rho_B}\{1 - \epsilon^{-2\rho_B}(1 + \rho_B)\}, \tag{40}$$

where ρ_B represents ρ_{B1}.

Hence the interaction energy of the charge distribution about the point B with a similar distribution about the point A is given by

$$J_1 = \frac{e^2}{a_0}\int \frac{u_A^2(1)}{\rho_B}\{1 - e^{-2\rho_B}(1 + \rho_B)\}dv_1$$

$$= \frac{e^2}{a_0}\cdot\frac{1}{\pi}\int \frac{\epsilon^{-2\rho_A}}{\rho_B}\{1 - \epsilon^{-2\rho_B}(1 + \rho_B)\}dv_1.$$

Introducing the confocal elliptic coördinates defined by equations (37*a*) and (37*b*), and integrating with respect to θ, the last equation assumes the form

$$J_1 = \frac{e^2}{a_0}\cdot\frac{D^2}{2}\int \frac{\epsilon^{-D(\lambda+\mu)}}{(\lambda - \mu)}(\lambda^2 - \mu^2)d\lambda d\mu \cdot$$

$$\left\{1 - \epsilon^{-D(\lambda-\mu)}\left(1 + \frac{\lambda D}{2} - \frac{\mu D}{2}\right)\right\}$$

$$= \frac{e^2}{a_0}\cdot\frac{D^2}{2}\left\{\int_1^\infty \epsilon^{-D\lambda}d\lambda\int_{-1}^1 \epsilon^{-D\mu}(\lambda + \mu)d\mu - \right.$$

$$\left.\int_1^\infty \epsilon^{-2D\lambda}d\lambda\int_{-1}^1 (\lambda + \mu)\left(1 + \frac{\lambda D}{2} - \frac{\mu D}{2}\right)d\mu\right\}$$

$$= \frac{e^2}{a_0}\left\{\frac{1}{D}\left(\int_D^\infty \epsilon^{-x}x dx\int_{-D}^D \epsilon^{-x}dx + \int_D^\infty \epsilon^{-x}dx\int_{-D}^D \epsilon^{-x}dx\right)\right.$$

$$\left. - \frac{1}{4}\int_{2D}^\infty \epsilon^{-x}x dx - \frac{1}{16}\int_{2D}^\infty \epsilon^{-x}x^2 dx + \frac{D^2}{12}\int_{2D}^\infty \epsilon^{-x}dx\right\}$$

$$= \frac{e^2}{a_0}\left\{\frac{1}{D} - \epsilon^{-2D}\left(\frac{1}{D} + \frac{11}{8} + \frac{3D}{4} + \frac{D^2}{6}\right)\right\}. \tag{41}$$

In the case of the four negative terms on the right-hand side of equation (33), which involve respectively the four reciprocal distances ρ_{A1}^{-1}, ρ_{B2}^{-1}, ρ_{A2}^{-1}, and ρ_{B1}^{-1}, it follows from the equivalence of the

electrons that each of these terms represents the attractive energy due to interaction of an electron charge distribution about one of the nuclei, say A, with the positive charge at the other nucleus B. Denoting each of these four terms by J_2, it follows from equation (40) and the relation $R = a_0 D$ that

$$4J_2 = \frac{4e^2}{a_0 D}\{1 - \epsilon^{-2D}(1 + D)\}. \tag{42}$$

Consequently,

$$E_{11} = \frac{e^2}{R} + J_1 - 2J_2 \tag{43}$$

$$= \frac{e^2}{a_0} \cdot \frac{\epsilon^{-2D}}{D}\left\{1 + \frac{5D}{8} - \frac{3D^2}{4} - \frac{D^3}{6}\right\}. \tag{44}$$

From the plot in Fig. 56 of E_{11} against D it will be observed that for $D > 1.4$ approximately, the expression has a negative value. This signifies that, for distances beyond $D = 1.4$, the Coulomb *attractive energy* between the nuclei and the negative charge distributions exceeds the total repulsive energy which exists between the two nuclei and also between the two electronic charge distributions. The Coulomb *forces* of attraction and of repulsion are equal at the distance $D = 1.9$, approximately, for which $\partial E_{11}/\partial D = 0$.

We now have to evaluate the different integrals which occur in equation (34) for the energy E_{12}. Because of the equivalence of the electrons, it follows that the four negative terms must each be equal to the same integral. Let us consider the integral

$$\int V_1^{\mathrm{A}} \phi_1 \phi_2 dv_1 dv_2 = \frac{e^2}{a_0} \int u_{\mathrm{A}}(2) u_{\mathrm{B}}(2) dv_2 \int \frac{u_{\mathrm{A}}(1) u_{\mathrm{B}}(1) dv_1}{\rho_{\mathrm{A}1}} = SK_2,$$

where

$$K_2 = \frac{e^2}{a_0 \pi} \int \frac{\epsilon^{-(\rho_{\mathrm{A}1} + \rho_{\mathrm{B}1})}}{\rho_{\mathrm{A}1}} dv_1. \tag{45a}$$

Using the transformation to confocal elliptic coördinates defined by equations (37) and (38), it follows that

$$K_2 = \frac{e^2}{a_0} \cdot \frac{D^2}{2} \int_1^\infty \epsilon^{-D\lambda} d\lambda \int_{-1}^1 (\lambda - \mu) d\mu$$

$$= \frac{e^2}{a_0} \cdot \epsilon^{-D}(1 + D), \tag{45b}$$

and therefore,

$$SK_2 = \frac{e^2}{a_0} \cdot \epsilon^{-2D}\left(1 + 2D + \frac{4D^2}{3} + \frac{D^3}{3}\right). \tag{45c}$$

The first integral on the right-hand side of equation (34) is

$$\int V_0\phi_1\phi_2 dv_1 dv_2 = \frac{e^2}{R} \cdot S^2 + K_1,$$

where

$$K_1 = \frac{e^2}{a_0}\int \phi_1\phi_2\left(\frac{1}{\rho_{12}}\right) dv_1 dv_2. \tag{46}$$

Consequently, we obtain the result

$$E_{12} = e^2 \cdot \frac{S^2}{R} + K_1 - 2SK_2. \tag{47}$$

Heitler and London did not evaluate the integral in equation (46), but concluded that

$$K_1 < \frac{e^2}{a_0} \cdot \frac{5S^2}{8}.$$

However, Y. Sugiura[11] showed, as a result of a lengthy calculation, that the integral could be represented by the relation

$$K_1 = \frac{e^2}{5a_0}\left[-\epsilon^{-2D}\left(-\frac{25}{8} + \frac{23D}{4} + 3D^2 + \frac{D^3}{3}\right) + \right.$$

$$\left. \frac{6}{D}\left\{S^2(C + \ln D) + S_1^2 Ei(-4D) - 2SS_1 Ei(-2D)\right\}\right], \tag{48}$$

where C = Euler's constant = 0.5772,

$$S_1 = \epsilon^D\left(1 - D + \frac{D^2}{3}\right),$$

ln = natural logarithm,

$$Ei(x) = \text{integral logarithm}^{12} = \int_{\infty}^{-x} \frac{\epsilon^{-u}}{u}\, du,$$

$$Ei(-x) = C + \frac{1}{4}\ln x^4 - x + \frac{1}{2}\cdot\frac{x^2}{2!} - \frac{1}{3}\cdot\frac{x^3}{3!} + \ldots (x \leqslant 17).$$

[11] Sugiura, *Z. Physik*, **45**, 484 (1927).

[12] Values of this function are given in "Funktionentafeln," by E. Jahnke and F. Emde, B. G. Teubner, Berlin, 1928, pp. 19–22. Also see Appendix III.

As a result of his calculation, Sugiura derived a minimum value for the interaction energy η_α of -3.2 v.e. for $D = 1.4$ ($R = 0.80$ Å), which is about 0.8 v.e. lower than that obtained by HL, but still above the observed value -4.72 v.e.[13]

12.5 The Variational Method. The HL method yields only an approximate value for the energy of formation of H_2 from the atoms; more accurate results have been obtained in the solution of this problem, as in that of the helium atom, by the application of the variational method. As a first illustration of this, it is interesting to observe how much more readily the eigenvalues may be derived in this manner as compared with the rather tedious calculation by HL in which the perturbation method was used.

As shown in the previous chapter, we can write the condition for the solution of the S. equation in the form

$$E = \frac{\int \phi H \phi d\tau}{\int \phi^2 d\tau} = \text{minimum}, \tag{49}$$

where H is defined by equation (12), and the value of E will be lower (more negative), the more nearly the function ϕ approaches the correct eigenfunction for the given state.

As a first approximation let us assume that the zero-order function is represented by the Heitler-London function ϕ^0 defined by the relation

$$\phi^0 = a\phi_1 + b\phi_2, \tag{50}$$

where

$$\phi_1 = u_A(1)u_B(2),$$

$$\phi_2 = u_A(2)u_B(1),$$

and $u_A(1)$, etc., are defined by equations (1) and (2), while a and b are parameters to be determined from the conditions

$$\frac{\partial E}{\partial a} = \frac{\partial E}{\partial b} = 0.$$

Since ϕ_1 and ϕ_2 are orthonormalized functions; it follows that equation (49) assumes the form

$$(a^2 + b^2 + 2abS^2)E = a^2H_{11} + ab(H_{12} + H_{21}) + b^2H_{22}, \tag{51}$$

[13] The energy of formation of H_2 from the atoms is, of course, $+4.72$ v.e. This value is greater than the observed energy of dissociation of H_2 because the latter does not include the " zero point " energy of vibration of the nuclei, as will be discussed more fully in the following chapter.

where S^2 is defined by equation (19), and

$$\left.\begin{aligned} H_{11} &= \int \phi_1 H \phi_1 d\tau \\ H_{12} &= \int \phi_1 H \phi_2 d\tau \\ H_{21} &= \int \phi_2 H \phi_1 d\tau \\ H_{22} &= \int \phi_2 H \phi_2 d\tau \end{aligned}\right\}. \tag{52}$$

Because of the equivalence of the two electrons, it is evident that

$$H_{11} = H_{22} \quad \text{and} \quad H_{12} = H_{21}.$$

Hence, equation (51) becomes

$$(a^2 + b^2 + 2abS^2)E = (a^2 + b^2)H_{11} + 2abH_{12}.$$

Differentiating with respect to a and b separately, we derive two relations which must be satisfied. These are

$$(a^2 + b^2 + 2abS^2)\frac{\partial E}{\partial a} = 2a(H_{11} - E) + 2b(H_{12} - S^2E) = 0,$$

and

$$(a^2 + b^2 + 2abS^2)\frac{\partial E}{\partial b} = 2a(H_{12} - S^2E) + 2b(H_{11} - E) = 0.$$

Consequently, we derive the determinantal relation

$$\begin{vmatrix} H_{11} - E, & H_{12} - S^2E \\ H_{12} - S^2E, & H_{11} - E \end{vmatrix} = 0,$$

that is, the secular equation,

$$(H_{11} - E)^2 - (H_{12} - S^2E)^2 = 0. \tag{53}$$

The roots of this equation are, evidently,

$$E_S = \frac{H_{11} + H_{12}}{1 + S^2}, \tag{54}$$

and

$$E_A = \frac{H_{11} - H_{12}}{1 - S^2}. \tag{55}$$

From equations (12) and (26), it follows that

$$H_{11} = \int \phi_1^2(2E_0 + V_0 - V_1^B - V_2^A)dv_1dv_2$$

$$= 2E_0 + \frac{e^2}{R} + e^2 \int \left(\frac{1}{r_{12}}\right)\phi_1^2 dv_1dv_2$$

$$-e^2 \int \left(\frac{1}{r_{B1}} + \frac{1}{r_{A2}}\right)\phi_1^2 dv_1dv_2$$

$$= 2E_0 + \frac{e^2}{R} + J_1 - 2J_2, \tag{56}$$

where J_1 and J_2 are expressions defined by equations (41) and (42), respectively. Comparing with equation (43), it is seen that

$$H_{11} = 2E_0 + E_{11}. \tag{57}$$

In the same manner it is shown that

$$H_{12} = \int \phi_1\phi_2\{2E_0 + V_0 - V_1^A - V_2^B\}dv_1dv_2$$

$$= S^2\left(2E_0 + \frac{e^2}{R}\right) + e^2 \int \left(\frac{1}{r_{12}}\right)\phi_1\phi_2 dv_1dv_2$$

$$-e^2 \int \left(\frac{1}{r_{A1}} + \frac{1}{r_{B2}}\right)\phi_1\phi_2 dv_1dv_2$$

$$= S^2\left(2E_0 + \frac{e^2}{R}\right) + K_1 - 2SK_2, \tag{58}$$

where K_2 and K_1 are the expressions defined by equations (45a) and (46), respectively. Comparing with equation (47), it follows that

$$H_{12} = 2E_0S^2 + E_{12}. \tag{59}$$

Hence,

$$E_S - 2E_0 = \frac{E_{11} + E_{12}}{1 + S^2} = \eta_\alpha \text{ [see equation (32)]}$$

$$= \frac{e^2}{R} + \frac{J_1 - 2J_2 + K_1 - 2SK_2}{1 + S^2}, \tag{60}$$

while

$$E_A - 2E_0 = \frac{E_{11} - E_{12}}{1 - S^2} = \eta_\beta \text{ [see equation (35)]}$$

$$= \frac{e^2}{R} + \frac{J_1 - 2J_2 - K_1 + 2SK_2}{1 - S^2}. \tag{61}$$

These results are identical with those derived in the previous section.

12.6 Method of Molecular Orbitals. The fact that the Heitler-London method leads to a binding energy of only 3.2 v.e. as compared with the observed value of 4.72 v.e. shows that the simple assumption made by Heitler and London for the form of ϕ_0 requires considerable modification.

In seeking for an explanation of the discrepancy between calculated and observed values, it is evident that one cause is the neglect of the mutual polarization of the atoms. As mentioned in the following section, N. Rosen has shown that, when this factor is taken into account, the value deduced for the binding energy is 4.02 v.e.

Further consideration shows that in the Heitler-London method two very important effects have been neglected.[14] In the first place no allowance has been made for the probability of the simultaneous occurrence of both electrons near the same nucleus. This would be indicated by the presence in the eigenfunction of ionic terms. "When they are included," as Penney remarks, "they must be introduced symmetrically into the wave function in order that there is just as much probability of finding both electrons on the one proton as the other. If this were not done, the assumed wave functions would attribute a permanent dipole moment to the hydrogen molecule."

In the second place the Heitler-London method fails to take into account to a sufficient extent the fact that the two electrons will tend to avoid each other, so that the probability of the occurrence of one electron at any point in the configuration space must be a function of r_{12}. The additional energy of binding which results from this effect is the *correlation energy*, and while it is included to some extent in the exchange term E_{12}, the actual calculation of this term presents considerable difficulty.

The difficulties arise, of course, from the fact that in the Heitler-London method the zero-order eigenfunctions for the molecule are built up from *atomic orbitals*, that is, single-electron wave functions which

[14] W. G. Penney, "The Quantum Theory of Valency," p. 18; S. Dushman and F. Seitz, *loc. cit.*

describe the behavior of an electron in the field of only one proton. It is obvious, however, that, when the atoms combine to form a hydrogen molecule, these atomic orbitals will be profoundly modified, and it might therefore appear more logical to start with a system consisting of two nuclei at a given distance, to which the two electrons are added in succession. That is, we can regard H_2 as formed from the ion H_2^+ by the addition of an electron.

This method which has been developed by F. Hund, R. S. Mulliken,[15] and others is known as that of *molecular orbitals*. In this method the wave function expresses the motion of each electron in the potential field resulting from all the nuclei and the other electrons present in the molecule. The point of view is therefore analogous to that of Hartree in the case of atomic systems.

A comparison of the results obtained by the two methods has been made by Van Vleck and Sherman, and their conclusion is as follows:[16]

> It is hard to say categorically whether the method of molecular orbitals or the HL method is the better. The latter undoubtedly is much preferable at very large distances of separation of the atoms, for then the continual transfer of electronic charge from one atom to another demanded by the ionic terms surely scarcely occurs at all. On the other hand, at small distances, the HL method probably represents excessive fear of the r_{12} effect, and the factorization into n one-electron problems presupposed by the method of molecular orbitals may be quite a good approximation.

The reference to "ionic terms" requires further amplification, and, in order to illustrate the significance of this point, we shall consider the application of the two methods to a diatomic molecule AB in which each atom contains only one electron which is effective for interaction.

In the molecular orbital method, we represent the eigenfunction for the molecule by

$$\Psi = \psi_1(x_1y_1z_1) \cdot \psi_2(x_2y_2z_2), \tag{62}$$

where a capital Ψ is used, as Van Vleck and Sherman have suggested, to denote the wave function for the whole system, and small ψ for the function of each electron moving in the field of the two nuclei and the other electron.

Let ϕ_A denote the wave function for the motion of one electron in the field of the nucleus A (so that ϕ_A is an atomic orbital), and similarly for ϕ_B.

[15] R. S. Mulliken, *Rev. Modern Phys.*, **4**, 1 (1932). See also J. H. Van Vleck and A. Sherman, *ibid.*, **7**, 167 (1935) and W. G. Penney, *op. cit.*, Chapter III, for a discussion of this method.

[16] Van Vleck and Sherman, *op. cit.*, p. 171.

Then equation (62) becomes

$$\begin{aligned}\Psi &= [\alpha\phi_A(1) + \beta\phi_B(1)]\,[\alpha\phi_A(2) + \beta\phi_B(2)] \\ &= \alpha^2\phi_A(1)\phi_A(2) + \beta^2\phi_B(1)\phi_B(2) \\ &\quad + \alpha\beta[\phi_A(1)\phi_B(2) + \phi_A(2)\phi_B(1)].\end{aligned} \tag{63}$$

It will be noted that a term such as $\phi_A(1)\phi_A(2)$ implies that both electrons are associated with nucleus A, while $\phi_B(1)\phi_B(2)$ means that both electrons are associated with B. Therefore, the right-hand side of (63) represents a condition which may be described as follows:

(1) The system has a probability of being in either or both (depending upon the ratio α/β) of the *ionic states* A: B or A :B. These correspond to the designations A^-B^+ and A^+B^-, respectively.

(2) There is a probability, defined by $\alpha^2\beta^2$, that the molecule will exist in the *homopolar* state A : B, in which the bond is *non-polar*.

The HL argument takes into account only the second possibility and neglects completely the possible occurrence of ionic states.

Obviously the actual wave function should, in general, be represented by a combination of ionic and homopolar functions, and for purposes of further discussion we may write equation (63) in the more compact form

$$\Psi = a\psi_i + c\psi_{HL}, \tag{64}$$

where ψ_i designates the wave function for the ionic states, ψ_{HL} that for the homopolar state (HL form of function), and it is necessary that

$$a^2 + c^2 = 1.$$

From this point of view, the distinction between ionic and homopolar compounds is not nearly so sharp as we ordinarily assume. As a in equation (64) varies from 1 to 0, we pass from the completely ionic to completely homopolar type of molecule. In this connection the comments of Van Vleck and Sherman are very pertinent. They write:

> There are elements of truth in the old-fashioned chemistry that HCl has the structure H^+Cl^-, as the true wave function of HCl is expressible as a linear combination of various idealized types, and certainly H^+Cl^- must be given some representation. . . . One great service of quantum mechanics is to show very explicitly that *all gradations of polarity are possible*, so that in a certain sense it is meaningless to talk of such *idealizations* as homopolar bond, heteropolar bond, covalent bond, dative bond, etc.[17]

The reader will find further interesting remarks on this topic in Pauling's paper,[18] " The transition from one extreme bond type to

[17] Van Vleck and Sherman, *op. cit.*, p. 171. The author is responsible for the italicized parts.

[18] Pauling, *J. Am. Chem. Soc.*, **54**, 988 (1932). See also the discussion by L. Pauling and E. B. Wilson, Jr., " Introduction to Quantum Mechanics," p. 345.

another," in which he uses as illustrations of such transitions from ionic to homopolar molecules the alkali and hydrogen halides.

12.7 The Dissociation Energy of H_2. These considerations have been found to be important in the calculation of the energy of formation of H_2 from the atoms. (This will be designated by D_e.) For a symmetrical molecule, such as H_2, we usually regard the value of a in equation (64) as insignificant. However, S. Weinbaum[19] has shown that it is possible, by using a function of the form given in this equation, to obtain a more accurate value of D_e. He assumed

$$\Psi = [\phi_A(1)\phi_A(2) + \phi_B(1)\phi_B(2)] + c[\phi_A(1)\phi_B(2) + \phi_A(2)\phi_B(1)],$$

$$\phi_A(1) = \sqrt{\frac{Z^3}{\pi}}\, \epsilon^{-Z\rho_{A1}},$$

and similar expressions for the other hydrogen-like functions. This expression for Ψ was then used in the fundamental relation (49) to calculate a value of E, subject to the conditions $\partial E/\partial c = 0$ and $\partial E/\partial Z = 0$.

The maximum value for D_e obtained in this manner was 4.00 v.e., with the values $c = 0.256$ and the effective nuclear charge $Z = 1.193$.

E. A. Hylleraas[20] also used the method of molecular orbitals, starting with wave functions which are solutions for the ionized hydrogen molecule H_2^+. The value thus derived for D_e was 3.6 v.e.

Passing now to the consideration of other variational treatments in which atomic orbitals have been used, we find that S. C. Wang[21] used a modified HL function of the form

$$\Psi = C\{\epsilon^{Z(r_{A1}+r_{B2})} + \epsilon^{Z(r_{A2}+r_{B1})}\},$$

where C is a normalization constant, and Z, as before, represents an effective nuclear charge which was determined by solving equation (49), subject to the condition $\partial E/\partial Z = 0$.

A more elaborate expression for Ψ was utilized by N. Rosen.[22] He took into account the distortion in the charge distribution in each atom which must occur when the two atoms are brought together (the polarization effect). He writes:

> The simplest way to represent this distortion is to consider the radius of the atom to change with the distance from the other atom. This is effectively what Wang did in his calculations, and it led to a definite improvement in the energy value. However, since the perturbations involved are not spherically symmetrical this can-

[19] S. Weinbaum, *J. Chem. Phys.*, **1**, 593 (1933).
[20] E. A. Hylleraas, *Z. Physik*, **71**, 739 (1931).
[21] S. C. Wang, *Phys. Rev.*, **31**, 579 (1928).
[22] N. Rosen, *ibid.*, **38**, 2099 (1931).

not be a very good approximation to the true state of affairs, and the next improvement that suggests itself is to introduce a change in the wave function that will depend on the direction with respect to the molecular axis and will be greatest in the direction of the latter. Since the interactions can be thought of roughly as being along this axis, it seems likely that the electron cloud tends to bulge out in the direction of the second atom.

In accordance with this idea, Rosen used a modification of Wang's " trial function" as defined in the previous equation, of the form

$$\Psi = C[\epsilon^{-\alpha(r_{A1}+r_{B2})}(1 + \beta r_{A1}\cos\theta_{A1})\,(1 + \beta r_{B2}\cos\theta_{B2}) + \epsilon^{-\alpha(r_{B1}+r_{A2})}(1 + \beta r_{B1}\cos\theta_{B1})\,(1 + \beta r_{A2}\cos\theta_{A2})], \tag{65}$$

where θ_{A1} is the angle between r_{A1} and R (the line joining the nuclei), β is a variable parameter, and C is the normalization constant. Substituting this function in the variation equation (49), the effect was first determined of varying the effective nuclear charge α, assuming $\beta = 0$. The value of α thus obtained which corresponded to a minimum value of E was then substituted in equation (65) and used to calculate a minimum value of E for variations in β. The final result led to a value for D_e of 4.02 v.e., with the corresponding values of the two parameters $\alpha = 1.19$, and $\beta = 0.10$ for the equilibrium distance $R = 1.416a_0$.

The most accurate solution is that of H. M. James and A. S. Coolidge.[23] The method used by them is similar to that of Hylleraas for helium (see Chapter XI). They introduced the four elliptic coördinates

$$\lambda_1 = \frac{(r_{A1} + r_{B1})}{R}, \qquad \mu_1 = \frac{(r_{A1} - r_{B1})}{R},$$

$$\lambda_2 = \frac{(r_{A2} + r_{B2})}{R}, \qquad \mu_2 = \frac{(r_{A2} - r_{B2})}{R},$$

and the variable $\rho = 2r_{12}/R$, thus taking into account the electron correlations discussed in the previous section. As trial function they selected the *molecular orbital* defined by the series

$$\phi = \frac{1}{2\pi}\epsilon^{-\delta(\lambda_1+\lambda_2)}\sum C_{mnjkp}(\lambda_1^m\lambda_2^n\mu_1^j\mu_2^k\rho^p + \lambda_1^n\lambda_2^m\mu_1^k\mu_2^j\rho^p), \tag{66}$$

where the summation extends over all positive values of the exponents (including zero), " subject to the restriction required by nuclear symmetry that $j + k$ must be even, and taking as many terms as shall prove necessary to give an acceptable approximation for the energy."

[23] H. M. James and A. S. Coolidge, *J. Chem. Phys.*, **1**, 825 (1933).

As a first approximation only the exponential part was used, that is,

$$\phi = \frac{1}{2\pi}\epsilon^{-\delta(\lambda_1+\lambda_2)} = \frac{1}{2\pi}\epsilon^{-\frac{\delta}{R}(r_{A1}+r_{B1}+r_{A2}+r_{B2})}. \tag{67}$$

For each value of R there is a best value of δ. For the value $R = 1.40$ atomic units (the value deduced from observations on band spectra) the lowest values of $E = \int \phi H \phi d\tau$ leads to a binding energy of 2.56 v.e., which is comparable with that obtained by the Heitler-London function. As James and Coolidge remark, " In the energy thus calculated, there is nothing resembling the 'exchange integrals' of the HL treatment; this raises the question whether the importance of the exchange terms, frequently assumed to represent the essential nature and magnitude of chemical binding, may not have been overemphasized."

The fact, thus pointed out, that a form of molecular orbital function may be chosen that does not lead to integrals of the type H_{12} [see equations (52)] is of extreme significance for the physical interpretation of the Heitler-London calculation.

The next approximation made by James and Coolidge is the use of a series for ϕ in which the exponent p in equation (66) is put equal to zero. This means that the tendency of the electrons to avoid each other is completely neglected. Under these conditions the maximum value obtainable for the binding energy for $R = 1.40$ is about 4.27 v.e., which leaves about 0.5 v.e. to be accounted for by electron correlations.

The simplest expression in which ρ occurs has the form

$$\phi = \frac{\epsilon^{-\delta(\lambda_1+\lambda_2)}}{2\pi}\{C_0 + C_1(\mu_1^2 + \mu_2^2) + C_2\mu_1\mu_2 + C_3(\lambda_1 + \lambda_2) + C_4\rho\},$$

which involves five terms. The exact method used for the determination of the coefficients is described in the original publication. The value deduced for the binding energy by the use of this expression for ϕ is 4.507 v.e. for $R = 1.40$, with $\delta = 0.75$. Actually, calculations were carried out with as many as 13 terms in the series. Table 1 taken from the paper by James and Coolidge gives a comparison between their results and those of previous investigators.

It is of interest to compare these calculated values of D_e with the results of observations on the energy of dissociation of H_2, which we shall designate by D_0. Direct thermochemical determination by F. R. Bichowsky and L. C. Copeland[24] gave the value $D_0 = 4.55 \pm 0.15$ v.e., while H. Beutler[25] has deduced, from the observations on the

[24] F. R. Bichowsky and L. C. Copeland, *J. Am. Chem. Soc.*, **50**, 1315 (1928).

[25] H. Beutler, *Z. physik. Chem.*, **B29**, 315 (1935).

vibrational energy levels of H_2 in the normal state, the value $D_0 = 4.45$ ($= 102{,}700$ cal./mole). Since the energy minimum ($-D_e$), as calculated by the methods of quantum mechanics, includes the zero-point energy 0.27 v.e. ($=\frac{1}{2}h\nu_0$ as shown in the following chapter), the observed value of D_e is $4.45 + 0.27 = 4.72$ v.e. ($= 108{,}900$ cal./mole), which is in extremely satisfactory agreement with that deduced by James and Coolidge.

TABLE 1

Function	D_e (*electron volts*)	R/a_0
1 term	2.56	1.40
5 terms	4.507	1.40
11 terms	4.685	1.40
13 terms	4.698	1.40
Without r_{12}	4.27	1.40
Heitler-London	2.9	1.40
Sugiura	3.2	1.51
Wang	3.76	1.42
Rosen	4.02	1.416
Observed	4.72	1.40

SUPPLEMENTARY NOTE 1

TRANSFORMATION FROM CARTESIAN TO CONFOCAL ELLIPTIC COÖRDINATES

The consideration of Figs. 61*a* and 61*b* will illustrate the physical significance of this coördinate system.

If $2a$ designates the distance between two points regarded as foci of a system of ellipses and hyperbolas, then the major axis of any one ellipse is given by

$$CD = \frac{2a}{e_1},$$

where e_1 ($\leqslant 1$) is the eccentricity of the given ellipse.

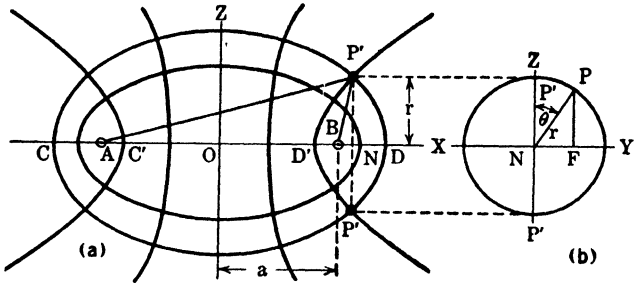


FIG. 61. Confocal elliptic coördinates.

Similarly, for any confocal hyperbola, the major axis is given by

$$C'D' = \frac{2a}{e_2},$$

where e_2 ($\geqslant 1$) is the eccentricity of the hyperbola.

For any point P', at which these two curves intersect,

$$AP' + BP' = r_A + r_B = \frac{2a}{e_1} = 2a\lambda, \tag{i}$$

and

$$AP' - BP' = r_A - r_B = \frac{2a}{e_2} = 2a\mu, \tag{ii}$$

where $\lambda = 1/e_1$ and $\mu = 1/e_2$.

Thus, each ellipse of the system of confocal conics has a definite value of λ, and similarly, each confocal hyperbola has a definite value of μ, so that a point P may be designated by specifying the values of λ and μ for the two confocal conics which intersect there.

If now we rotate these curves about AB as axis of revolution, there are obtained a series of confocal conicoids which in this case are known as *prolate spheroids.* The ellipsoids and hyperboloids of revolution will intersect along circles such as that indicated in Fig. 61*b*, and hence it is necessary to introduce a third coördinate variable, the angle θ, which specifies the position of P on the circumference of the circle, with respect to a fixed plane (indicated by the line $P'NP'$) which passes through the axis AB.

It is necessary now to derive the relations between the variables λ, μ, θ and the rectangular variables x, y, z. Assuming that the x-axis coincides with the axis of revolution AB, that the z-axis is in the plane of the figure and passes through the origin O, in the midpoint of AB, while the y-axis projects at right angles to the plane containing Oz and Ox, we can set down the following relations which are derived in any textbook on solid geometry.

$$\frac{x^2}{\lambda^2} + \frac{y^2 + z^2}{\lambda^2\left(1 - \frac{1}{\lambda^2}\right)} = a^2. \tag{iii}$$

$$\frac{x^2}{\mu^2} - \frac{y^2 + z^2}{\mu^2\left(1 - \frac{1}{\mu^2}\right)} = a^2. \tag{iv}$$

At the point P the same values of x, y, z must satisfy both of these equations. Hence,

$$r^2 = y^2 + z^2 = a^2(\lambda^2 - 1)\,(1 - \mu^2).$$

This defines the radius $r(= NP)$ of the circle along which the two surfaces intersect, and consequently, as is evident from Fig. 61*b*,

$$y = a\sqrt{(\lambda^2 - 1)\,(1 - \mu^2)} \cdot \sin\theta; \tag{v}$$

$$z = a\sqrt{(\lambda^2 - 1)\,(1 - \mu^2)} \cdot \cos\theta. \tag{vi}$$

By substituting in either equation (iii) or (iv), it is readily shown that

$$x = a\lambda\mu. \tag{vii}$$

As in section 6.2, we now have to determine the three constants a_λ, a_μ, and a_θ, which connect the elements of volume as expressed in the two systems of coördinates, according to the relation

$$dxdydz = \sqrt{a_\lambda a_\mu a_\theta} \cdot d\lambda d\mu d\theta.$$

Since λ, μ, and θ form an *orthogonal* system of coördinates, it follows that

$$a_\lambda = \left(\frac{ds}{d\lambda}\right)^2 = \left(\frac{\partial x}{\partial \lambda}\right)^2 + \left(\frac{\partial y}{\partial \lambda}\right)^2 + \left(\frac{\partial z}{\partial \lambda}\right)^2; \tag{viii}$$

$$a_\mu = \left(\frac{ds}{d\mu}\right)^2 = \left(\frac{\partial x}{\partial \mu}\right)^2 + \left(\frac{\partial y}{\partial \mu}\right)^2 + \left(\frac{\partial z}{\partial \mu}\right)^2; \tag{ix}$$

$$a_\theta = \left(\frac{ds}{d\theta}\right)^2 = \left(\frac{\partial x}{\partial \theta}\right)^2 + \left(\frac{\partial y}{\partial \theta}\right)^2 + \left(\frac{\partial z}{\partial \theta}\right)^2; \tag{x}$$

where $(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2$.

These differential coefficients, as derived from equations (v), (vi), and (vii), are as follows.

$$\left.\begin{aligned} \frac{\partial x}{\partial \lambda} = \mu a; \frac{\partial y}{\partial \lambda} &= \lambda a \sin\theta \sqrt{\frac{1-\mu^2}{\lambda^2-1}} \\ \frac{\partial z}{\partial \lambda} &= \lambda a \cos\theta \sqrt{\frac{1-\mu^2}{\lambda^2-1}} \end{aligned}\right\}; \tag{xi}$$

$$\left.\begin{aligned} \frac{\partial x}{\partial \mu} = \lambda a; \frac{\partial y}{\partial \mu} &= -\mu a \sin\theta \sqrt{\frac{\lambda^2-1}{1-\mu^2}} \\ \frac{\partial z}{\partial \mu} &= -\mu a \cos\theta \sqrt{\frac{\lambda^2-1}{1-\mu^2}} \end{aligned}\right\}; \tag{xii}$$

$$\left.\begin{aligned} \frac{\partial x}{\partial \theta} = 0; \frac{\partial y}{\partial \theta} &= a\sqrt{(\lambda^2-1)(1-\mu^2)} \cdot \cos\theta \\ \frac{\partial z}{\partial \theta} &= -a\sqrt{(\lambda^2-1)(1-\mu^2)} \cdot \sin\theta \end{aligned}\right\}. \tag{xiii}$$

Hence,

$$a_\lambda = \frac{a^2(\lambda^2-\mu^2)}{(\lambda^2-1)}; \tag{xiv}$$

$$a_\mu = \frac{a^2(\lambda^2-\mu^2)}{(1-\mu^2)}; \tag{xv}$$

$$a_\theta = a^2(\lambda^2-1)(1-\mu^2); \tag{xvi}$$

and

$$dxdydz = a^3(\lambda^2-\mu^2)d\lambda d\mu d\theta. \tag{xvii}$$

The same result follows from the relation in terms of the Jacobian, which is of the form

$$dxdydz = \begin{vmatrix} \frac{\partial x}{\partial \lambda} & \frac{\partial y}{\partial \lambda} & \frac{\partial z}{\partial \lambda} \\ \frac{\partial x}{\partial \mu} & \frac{\partial y}{\partial \mu} & \frac{\partial z}{\partial \mu} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} & \frac{\partial z}{\partial \theta} \end{vmatrix} \cdot d\lambda d\mu d\theta.$$

As may be verified by direct substitution from equations (xi), (xii), and (xiii), the value of the determinant is found to be $a^3(\lambda^2 - \mu^2)$.

Applying the rules stated in section (6.2) (see also Appendix IV) for expressing the Laplacian operator in terms of the variables λ, μ, and θ, it is readily deduced that

$$\nabla^2 = \frac{1}{a^2(\lambda^2 - \mu^2)}\left[\frac{\partial}{\partial \lambda}\left\{(\lambda^2 - 1)\frac{\partial}{\partial \lambda}\right\} + \frac{\partial}{\partial \mu}\left\{(1 - \mu^2)\frac{\partial}{\partial \mu}\right\} + \frac{(\lambda^2 - \mu^2)}{(\lambda^2 - 1)(1 - \mu^2)} \cdot \frac{\partial^2}{\partial \theta^2}\right]. \qquad \text{(xviii)}$$

When the S. equation for a two-center system is expressed in terms of these variables it is usually much more convenient to separate the equation into three ordinary differential equations, and thus obtain a solution, than when other systems of variables are used.[26]

[26] See discussion by H. Bethe in "Handbuch der Physik," XXIV, Part 1, p. 530 *et seq.* A further discussion of the usefulness of elliptic coördinates in the treatment of the two-center problem will be found in the paper by W. G. Barber and H. R. Hassé, *Proc. Cambridge Phil. Soc.*, **31**, 564 (1935).

CHAPTER XIII

VIBRATIONAL AND ROTATIONAL STATES OF THE HYDROGEN MOLECULE

13.1 General Remarks.[1] The existence of band spectra, and the observations on the variation in specific heats of gases with temperature, lead to the conclusion that, in addition to the energy of excitation of electronic levels, the molecules also possess both vibrational and rotational energy due to the motions of the nuclei. Since the frequencies of these motions are small compared to those of the electrons, the electronic motion can adjust itself relatively instantaneously to the motion of the nuclei as if the latter were centers of force at rest.

It is therefore possible to consider the energy states arising from nuclear motions as superimposed upon the electronic states.

Let us consider a diatomic molecule, such as H_2. As shown in the previous chapter, the potential energy for any two atoms A and B, which combine to form a molecule AB, is a function of the internuclear distance r, which may be represented graphically by a curve such as that shown in Fig. 62. Let $U(r)$ designate this function. By definition, the force between the nuclei is given by

$$F(r) = -\frac{dU}{dr}.$$

At the value of $r = r_0$, for which U is a minimum, the force vanishes, while it is negative to the right of the minimum (corresponding to a net force of attraction) and positive to the left of the minimum (corresponding to a net force of repulsion). If we assume $U(r) = 0$ for $r = \infty$, then the energy at $r = r_0$ (which is negative) corresponds to the dissociation energy which is usually designated by $-D$.

For values of $U(r) > -D$, there are two values of r at which $U(r)$ has the same value. These correspond to mean points of equilibrium for the vibrational motion of the nuclei, and the total energy is given by $E_v = U(r) + D$. This energy is quantized; so that there exist a series of vibrational energy states, designated by the quantum numbers $v = 0, 1, 2, 3$, etc., between the values $U(r) = -D$ and $U(r) = 0$. Figure 62 shows such a series of vibrational energy states for the mole-

[1] See references at end of chapter.

cule H_2 in the normal ($1s'\Sigma_g^+$) state, and the potential energy curve derived in accordance with Morse's equation, as explained in the following section.

It is also observed that with any given vibrational state there are associated a series of rotational energy states, designated by the quantum numbers $K = 0, 1, 2$, etc. Thus, any line in the emission spectrum

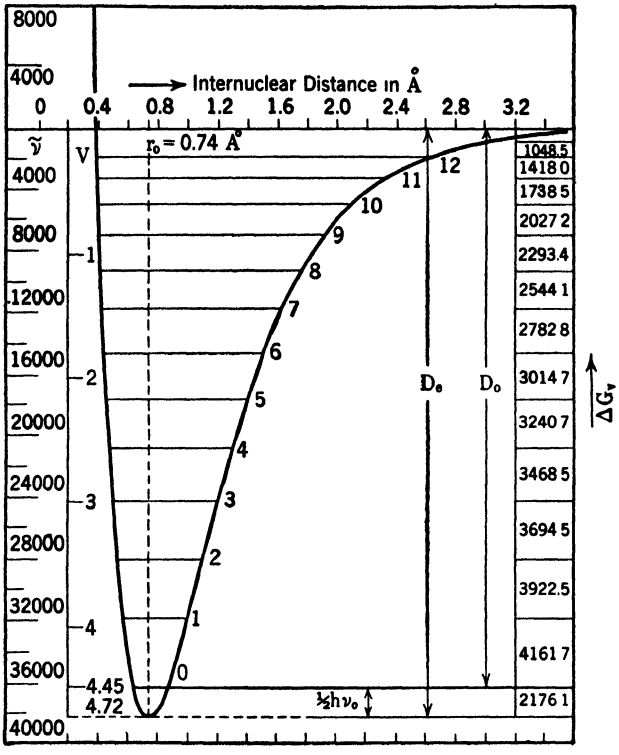


FIG. 62. Potential energy function and vibrational energy levels for normal state of H_2 molecule.

of a molecule may be represented as a transition from an upper rotational level K', associated with the upper vibrational level v' and with a higher electronic state, to a lower rotational level K'', which is associated with a vibrational level v'' and a lower electronic state. That is, if $\tilde{\nu}$ represents the wave number[2] of the emitted line,

$$\tilde{\nu} = \tilde{\nu}_e + G'_v - G''_v + F'(v,\ K) - F''(v,\ K), \qquad (1)$$

where $\tilde{\nu}_e$ denotes the wave-number difference between the lowest levels

[2] It is customary to give energy differences in terms of wave numbers. See Appendix II for conversion factors to other units.

associated with $v' = v'' = 0$ and $K' = K'' = 0$, G'_v corresponds to the vibrational energy, and $F'(v, K)$ to the rotational energy associated with the higher electronic state, and similarly for G''_v and $F''(v, K)$.

Thus Fig. 63 shows the lower vibrational energy levels for the molecule H_2 in the normal $(1s'\Sigma_g^+)$ and in the first excited $(2p'\Sigma_u^+)$ states.[3] The latter state results from the interaction of one electron in the normal (1s) state and another in the 2s state. The difference in energy

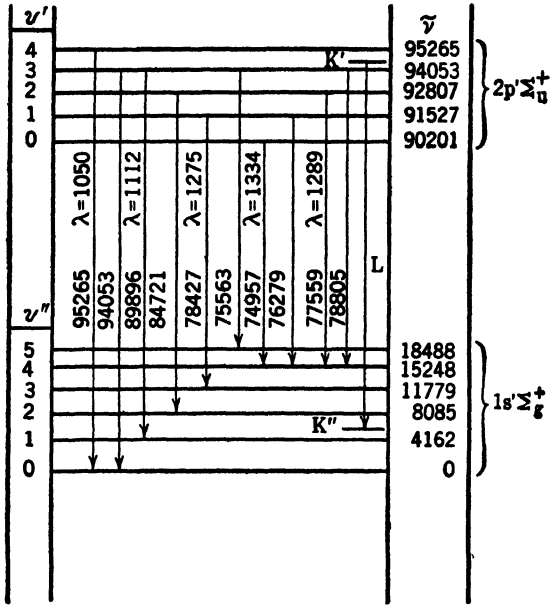


FIG. 63. Normal and first excited electronic energy levels for H_2 molecule and associated vibrational energy levels.

for the states $v = 0$, $K = 0$ is equal to 90,201 cm.$^{-1}$ (= 11.13 v.e.), and the spectrum emitted by transitions from the higher to the normal state lies in the ultraviolet region, as shown by the values of λ (in Ångström) which are attached to the different lines.

Figure 64 shows the rotational energy levels associated with the level, $v' = 3$, in the upper electronic state, and the level, $v'' = 1$, in the normal state of the H_2 molecule. The values of $\tilde{\nu}$ on the right-hand side are taken from the corresponding values in Fig. 63,[4] while the values

[3] The values of the energy levels for H_2 indicated in this and the other figures are taken from the extremely interesting paper by C. R. Jeppesen, *Phys. Rev.*, **44**, 165 (1933), "The Emission Spectrum of Molecular Hydrogen in the Extreme Ultraviolet."

[4] The difference between the values 4162 and 4157 for $\tilde{\nu}_0$ of the level $v'' = 1$ is of no significance in the present connection.

under $\Delta\tilde{\nu}$ give the total increase in rotational energy (in terms of wave numbers) as the quantum number K is increased from 0 to higher values. The transitions indicated are those corresponding to the line L in Fig. 63, where K' designates the upper levels and K'' the lower levels, respectively. It will be seen that the " head " of the band, corresponding to the transition from $K' = 0$ to $K'' = 0$, has the value $\tilde{\nu} = 89896$ (or $\lambda = 1112$ Å). This is shown in Fig. 63 as the third line from the left.

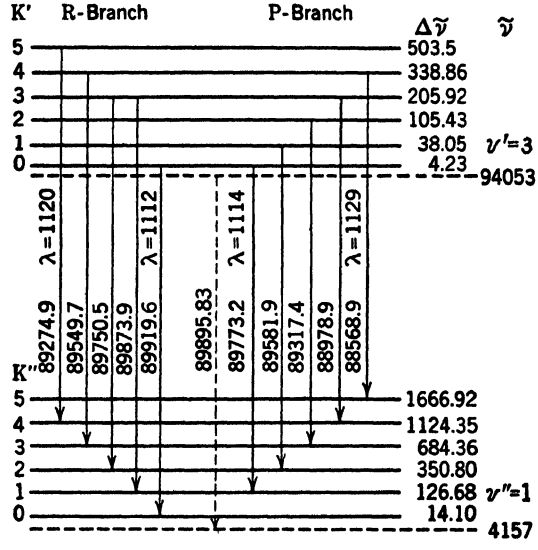


FIG. 64. Illustrating transitions between rotational energy levels associated with two different vibrational energy levels for H_2 molecule.

13.2 Vibrational Energy States. For the H_2 molecule in the normal state, the equilibrium internuclear distance is $r_0 = 0.74$ Å. As the molecule acquires increasing amounts of vibrational energy (due to increase in temperature of the gas), this internuclear distance increases. Assuming that the restoring force acting on each atom in a diatomic molecule AB is proportional to the displacement from the position of equilibrium for minimum value of $U(r)$ (Hooke's law), it is possible to derive relations for both the potential energy as a function of r and the frequency of vibration.[5]

We shall consider the general case of a diatomic molecule consisting of atoms A and B. Let μ_A and μ_B denote the masses of the atoms; let r denote the distance between the two nuclei for any given vibrational state, and let r_0 designate the value of r for $U(r) = -D$.

[5] R. de L. Kronig, " The Optical Basis of the Theory of Valency," p. 83.

Then, the equations of motion for the two nuclei are as follows (see Fig. 65):

$$\mu_A \ddot{x}_1 = k(r - r_0) = -k(x_1 - x_2),$$
$$\mu_B \ddot{x}_2 = -k(r - r_0) = k(x_1 - x_2).$$

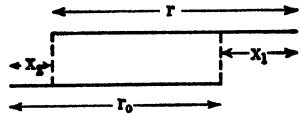


FIG. 65. Illustrating the vibrational motion of two nuclei.

Hence,

$$\ddot{x}_1 - \ddot{x}_2 = -k\left(\frac{1}{\mu_A} + \frac{1}{\mu_B}\right)(x_1 - x_2).$$

That is,

$$\ddot{r} = -\frac{k}{\mu}(r - r_0), \tag{2}$$

where $\mu = \dfrac{\mu_A \mu_B}{\mu_A + \mu_B}$ = " reduced " mass of molecule.

The solution of this equation is

$$r - r_0 = A \sin(2\pi\nu_0 t + \delta),$$

where A is the maximum amplitude, δ is the phase angle, and

$$\nu_0 = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}. \tag{3}$$

The potential energy function corresponding to equation (2) is evidently a parabola with vertex at $r = r_0$, the equation for which is

$$U(r) = -D + \frac{k}{2}(r - r_0)^2.$$

Substituting from equation (3), this can be written in the form

$$U(\rho) = -D + 2(\pi\omega_0)^2\mu\rho^2, \tag{4}$$

where $\rho = r - r_0$, and ν_0 is replaced by ω_0, in accordance with the notation used in the literature on band spectra.

This is the function which must be inserted in the S. equation for the system in order to determine the discrete energy states. This equation, which corresponds to equation (7.7), has the form

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{dS}{d\rho}\right) - \frac{K(K+1)S}{\rho^2} + \alpha^2\{E - U(\rho)\}S = 0, \tag{5}$$

where $S = S(\rho)$ is the " radial " function, $K = 0, 1, 2$, etc., and $\alpha^2 = 8\pi^2\mu/h^2$.

Put $S(\rho) = \frac{1}{\rho}\phi(\rho)$. Then equation (5) becomes

$$\frac{d^2\phi}{d\rho^2} - \frac{K(K+1)\phi}{\rho^2} + \alpha^2(E - U)\phi = 0. \tag{6}$$

If we substitute for $U(\rho)$ from equation (4), and let $K = 0$, that is, consider only the case in which there is *zero rotational energy*, and only vibrational energy, equation (6) becomes[6]

$$\frac{d^2\phi}{d\rho^2} + \alpha^2\{E + D - 2(\pi\omega_0)^2\mu\rho^2\}\phi = 0. \tag{7}$$

This equation is evidently similar to equation (5.5) for the linear harmonic oscillator, and has as eigenvalues the series defined by the relation

$$E_v = -D + h\omega_0(v + \tfrac{1}{2}), \tag{8}$$

where $v = 0, 1, 2$, etc., corresponds to the vibrational quantum number.

It follows that the vibrational energy levels should be equally spaced. Actually, the distance between successive energy levels decreases, in the case of homopolar diatomic molecules, with increase in v. That is, the atoms do not behave as simple harmonic oscillators. The motion is said to be of the *anharmonic type*, with the result that, to a first approximation, the vibrational energy must be represented by an expression of the form

$$E_v = -D + h\omega_0(v + \tfrac{1}{2}) - hx\omega_0(v + \tfrac{1}{2})^2, \tag{9}$$

where x is a constant for any given molecule.

For $E = 0$, that is, when dissociation occurs, it is evident that $dE_v/dv = 0$. Consequently,

$$\omega_0 - 2x\omega_0(v + \tfrac{1}{2}) = 0,$$

that is,

$$v + \frac{1}{2} = \frac{1}{2x},$$

[6] The solution discussed in the following section is that given by P. M. Morse, *Phys. Rev.*, **34**, 57 (1929); also in Condon and Morse, " Quantum Mechanics," Chapter V.

and

$$D = \frac{h\omega_0}{2x} - \frac{hx\omega_0}{4x^2} = \frac{h\omega_0}{4x}, \tag{10a}$$

or

$$\frac{D}{hc} = \frac{\omega_0^2}{c^2} \cdot \frac{c}{4\omega_0 x} \cdot \tag{10b}$$

The right-hand side of this equation is expressed in such a form as to give the value for the dissociation energy in terms of wave numbers.

Hence, to a first approximation,

$$E_v = -D + h\omega_0\left(v + \frac{1}{2}\right) - \frac{h^2\omega_0^2}{4D}\left(v + \frac{1}{2}\right)^2 \cdot \tag{11}$$

As shown by Morse, a potential energy function which is in agreement with this relation and with the requirement that E_v shall be a minimum for $\rho = 0$ may be chosen of the form

$$U(\rho) = -2De^{-\beta\rho} + De^{-2\beta\rho}, \tag{12}$$

where β is obtained as follows.

For small values of ρ, the last equation becomes

$$U(\rho) = -D + \beta^2\rho^2 D.$$

This evidently has a minimum value $U(\rho) = -D$ for $\rho = 0$, that is, the minimum occurs at $r = r_0$.

Comparing this with the equation (4) it follows that

$$\beta^2 D = 2\pi^2\mu\omega_0^2,$$

and substituting for D from equation (10a),

$$\beta = \sqrt{\frac{8\pi^2 M\omega_0 x}{h}} = 0.2454\sqrt{M\omega_0 x}, \tag{13}$$

where

$$M = \mu \times 6.064 \times 10^{23},$$

= reduced mass in terms of oxygen (atomic mass = 16).

Now C. R. Jeppesen[7] found that, for values of the vibrational quantum number from $v = -\frac{1}{2}$ to $v = 5\frac{1}{2}$, the difference in wave numbers between successive levels is given by the relation

$$\Delta G_v = 4417.19 - 262.63(v + \tfrac{1}{2}) + 9.34619(v + \tfrac{1}{2})^2 - 0.76(v + \tfrac{1}{2})^3. \tag{14}$$

The values of ΔG_v thus calculated are shown in the column on the

[7] C. R. Jeppesen, *loc. cit.*

right-hand side of Fig. 62. For $v = 1$, $\Delta G_v = 4161.70$, and it will be observed that the value of ΔG_v decreases with increase in v.

More recently H. Beutler[8] has been able to determine, from observations on the lines in the extreme ultraviolet spectrum, values of ΔG_v for $v = 12$, 13, and 14. If the values of ΔG_v are plotted as ordinates against $\sum_{v=0}^{v} \Delta G_v = \tilde{\nu}$, the total increase in wave number from $v = 0$ to v, the resulting curve (see Fig. 66) shows that $\Delta G_v = 0$ for $\tilde{\nu} = 36{,}116$ cm.$^{-1}$. Therefore, this value must correspond to the dissociation energy D_0 of H_2 in the normal state. In terms of energy units,

$$D_0 = \frac{36{,}116}{8106} = 4.455 \text{ v.e.}$$

$$= 4.455 \times 23{,}055 = 102{,}700 \text{ cal./mole.}$$

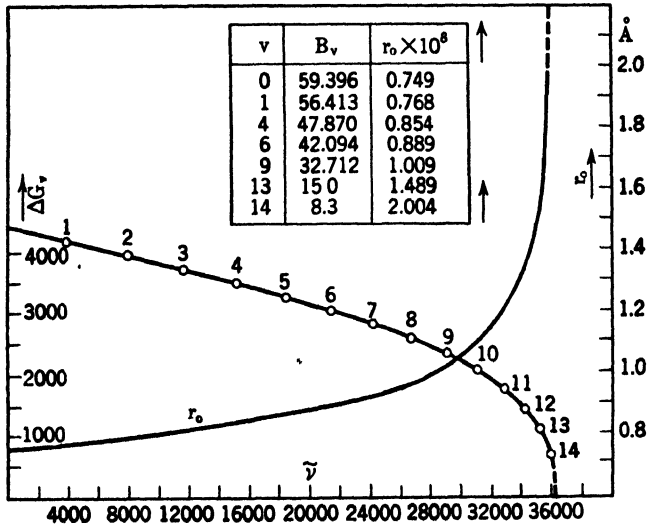


FIG. 66. Plot of ΔG_v and internuclear distance (r_0) as functions of the wave numbers of the vibrational levels for the normal state of H_2.

From equation (14) it follows that, for $v = -\frac{1}{4}$, $\Delta G_0 = 4352.2$.[9] This corresponds to twice the zero-point energy ($h\nu_0/2$), and therefore one-half this quantity must be added to D_0 to give D, the minimum value of the potential energy $U(\rho)$. Thus, $D = 36{,}116 + 2176 = 38{,}292$ cm.$^{-1}$ = 108,900 cal./mole.

[8] H. Beutler, *Z. physik. Chem.*, **B27**, 287 (1934); *ibid.*, **B29**, 315 (1935).

[9] According to Jeppesen the extrapolation to $v = -\frac{1}{4}$ gives the best approximation to the value of ΔG_0.

For H_2 ($M = 1.008$), the value of β according to equation (13) is 2.00. Hence the potential energy function for H_2, as derived from the spectroscopic vibrational energy levels, is given, *in terms of wave numbers*, by the relation

$$U(\rho) = 38{,}292(\epsilon^{-4\rho} - 2\epsilon^{-2\rho}). \tag{15}$$

The following table gives values of $U(\rho)$ for different values of $r = \rho + 0.74$, in Ångström units.

r(Å)	$U(\rho)$	r(Å)	$U(\rho)$
0.34	+18,997	1.14	−26,660
0 44	−12,265	1 34	19,578
0 54	29,062	1 54	13,906
0 64	36,354	1 74	9,664
0.74	38,291	1 94	6,633
0 84	37,075	2 34	3,061
0 94	34,123	2 74	1,390
1 04	30,532	3 74	192

Figure 62 was plotted from this data. The ordinates give the values of $U(\rho)$ both in terms of $\bar{\nu}$, the wave number, and in terms of V, the corresponding energy in electron volts. The horizontal lines give the vibrational energy levels corresponding to the different values $v = 0$ to $v = 12$, indicated on the curve. To distinguish D from D_0, the symbol D_e is customary (as mentioned in the previous chapter) for the former, which evidently has only a theoretical interest since D_0 *is the experimentally determined value.*

It will be observed that, according to Jeppesen's data, $2x\omega_0$ and ω_0, expressed in wave numbers, are equal to 262.63 and 4352.2 respectively. Hence, if we use equation (10*b*),

$$D = \frac{(4352.2)^2}{525.26} = 36{,}060 \text{ cm.}^{-1},$$

which is less than the value determined by direct extrapolation to $\Delta G_v = 0$. As first shown by R. T. Birge and H. Sponer,[10] the latter method is the more accurate one for the determination of D_0.

A potential energy curve similar to that for the normal state can also be plotted for any of the excited states for which sufficient data on the vibrational energy levels are available. According to Jeppesen, the values of ΔG_v for the $2p'\Sigma_u^+$ state are given by the relation

$$\Delta G_v = 1357.302 - 39.9307(v + \tfrac{1}{2}) + 1.218487(v + \tfrac{1}{2})^2 - 0.0638888(v + \tfrac{1}{2})^3,$$

[10] R. T. Birge and H. Sponer, *Phys. Rev.*, **28**, 259 (1926).

while the value corresponding to twice the zero-point energy is $\Delta G_0 = 1347.4$ cm.$^{-1}$. Since the value of $2x\omega_0$ is 39.931, the calculated value of D_e is $(1347.4)^2/79.862 = 22{,}860$ cm.$^{-1}$.

13.3 Rotational Energy Levels. Let us now consider the solution of equation (6) for the case $K \neq 0$. A solution of this equation has been given by C. L. Pekeris,[11] leading to an expression for the energy levels in which the effects of rotational and vibrational energies are combined. However, it is possible to indicate the significance of this more complicated expression without the formal mathematical derivation.

In equation (*6.25*) it was shown that in a rotating diatomic molecule with fixed moment of inertia I the rotational energy levels are given by the relation

$$E_r = \frac{h^2}{8\pi^2 I} K(K+1), \tag{16}$$

where K is the rotational quantum number, and I is derived by the relation

$$I = \mu r_0^2,$$

where μ is the reduced mass.

Equation (16) is usually written in the form

$$E_r = BK(K+1), \tag{17}$$

where

$$B = \frac{h^2}{8\pi^2 I}. \tag{18}$$

In the case of H_2, there are two atoms, each of mass $\mu_H = 1.662 \times 10^{-24}$ gm. In the state $v = 0$, the internuclear distance is 0.74×10^{-8} cm. Hence,

$$I = 4.55 \times 10^{-41} \text{ gm. cm.}^2,$$

$$B = 1.19 \times 10^{-14} \text{ gm. cm.}^2 \text{ sec.}^{-2},$$

and

$$B_v = \frac{B}{hc} = 60.65 \text{ cm.}^{-1}.$$

The latter is to be compared with the value 59.4 derived from the plot in Fig. 66. With increase in vibrational quantum number, B_v

[11] C. L. Pekeris, *Phys. Rev.*, **45**, 98 (1934). See also L. Pauling and E. B. Wilson, Jr., "Introduction to Quantum Mechanics," Chapter X, for detailed discussion of this case.

decreases, as shown in this figure, and r_0 increases in accordance with the relation

$$r_0^2 = 3.32 \times \frac{10^{-15}}{B_v}.$$

The notation B_v is used to indicate that the constant is measured in terms of wave numbers and also that the value is a function of the vibrational quantum number.

According to equation (16), the levels should be equally spaced, if I were a constant; but, as in the case of the vibrational levels, not only does I vary with v, but also, for a given value of v, the spacing between rotational levels varies with change in K. Thus, in the case of H_2 and many other molecules, the value of ΔF, the increase in wave number for $\Delta K = 1$, increases with increase in K, as is evident from Fig. 64. Consequently, the actual observations on rotational energy levels lead to the semi-empirical relation

$$\frac{E_r}{hc} = F = -\frac{1}{4}\cdot\frac{B}{hc} + \frac{B}{hc}\left(K+\frac{1}{2}\right)^2 + \frac{A}{hc}\left(K+\frac{1}{2}\right)^3 + \text{higher powers of } \left(K+\frac{1}{2}\right), \qquad (19)$$

where F is expressed in terms of wave numbers.

The first two terms on the right-hand side of this equation evidently correspond to the expression in equation (17), while the following terms are due to the effect of the vibrational energy. The expression deduced by Pekeris, which expresses the combined effects of rotational and vibrational energy, is:[12]

$$\frac{E_{K,v}}{hc} = \omega_0\left(v+\frac{1}{2}\right) - x\omega_0\left(v+\frac{1}{2}\right)^2 + B_v K(K+1) - \beta_0 K^2(K+1)^2 - \gamma_0\left(v+\frac{1}{2}\right)K(K+1),$$

where

$$x = \frac{h\omega_0 c}{4D} \ldots \text{[see equation (10}a\text{)]};$$

$$B_v = \frac{h}{8\pi^2 Ic} \ldots \text{[see equation (18)]};$$

[12] L. Pauling and E. B. Wilson, Jr., *op. cit.*, p. 274.

and β_0 and γ_0 are constants, the values of which depend on those of r_0, ω_0, μ, and D.

In Fig. 64 the values under $\Delta\tilde{\nu}$ give the difference in wave number of any given level with respect to the imaginary level designated by a dashed line, while $\tilde{\nu}$ gives the wave number of this lowest level with respect to the level $\tilde{\nu} = 0$ for the normal H_2 molecule in state $v = 0$, $K = 0$. It will be observed that the differences between successive levels in both sets increase with increase in K, and that these differences for any pair of values of K' (the upper values) are less than for the corresponding pair in the set K'' (lower levels).

Although it is possible to have transitions between any pair of vibrational levels, as shown in Fig. 63, this is not valid for rotational levels. According to the selection principle for transitions, only those transitions may occur for which $\Delta K = \pm 1$. The lines corresponding to the transition from K' to $K'' = K' + 1$ constitute the P-branch, while those corresponding to the transition from K' to $K'' = K' - 1$ constitute the R-branch. In some cases, for reasons which we need not discuss here, lines are obtained for which $K'' = K'$. These then constitute the Q-branch.

From equation (19) it is seen that the difference between the level $K = 0$ and the dashed level immediately below it, in Fig. 64, is equal to $(\frac{1}{4})B_v$. Hence for the state $v' = 1$, $B_v' = 4 \times 14.1 = 56.1$ cm.$^{-1}$, while for the level $v'' = 3$ in the upper electronic state, $B_v'' = 16.9$ cm.$^{-1}$. According to Jeppesen, the values of B_v for the different vibrational levels associated with the normal state are given by the relation

$$B_v = 60.8715 - 3.06709(v + \tfrac{1}{2}) + 0.068393(v + \tfrac{1}{2})^2 - 0.0065(v + \tfrac{1}{2})^3.$$

Figure 67, taken from the original publication, shows the variation in B_v with v for three different electronic states of the H_2 molecule.

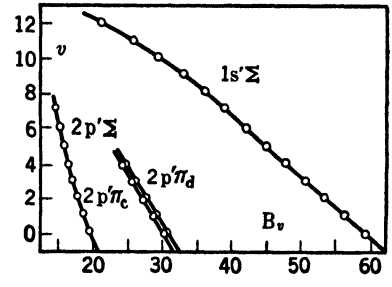


FIG. 67. Plot of B_v versus v for the normal ($1s^1\Sigma$) and two excited electronic states of H_2.

It follows from equation (18) and these plots that r *increases with increases in* v. That is, as the vibrational energy increases, the atoms tend to vibrate about positions of equilibrium which are increasingly farther apart, and when dissociation occurs, B_v vanishes and r_0 becomes infinitely great. In Fig. 66 there is shown a plot of r_0 versus $\tilde{\nu}$ which illustrates this conclusion graphically. It will be observed that, while the value of r_0 for $v = 0$ is 0.75×10^{-8}, this increases with increase in

v, and the rate of increase becomes greater as the stage of dissociation is approached, so that between $v = 13$ and $v = 14$, r_0 increases from 1.49 to 2.00×10^{-8} cm., and becomes infinite[13] between $v = 14$ and $v = 15$.

13.4 Orthohydrogen and Parahydrogen. One of the most brilliant achievements of quantum mechanics was the deduction by W. Heisenberg[14] and F. Hund[15] that hydrogen must occur in two forms analogous to ortho- and para-helium. These conclusions were supported by the observations on the relative intensities of rotational lines in the band spectrum of H_2. D. M. Dennison[16] applied the same suggestion to interpret the observed decrease in rotational specific heat of hydrogen at very low temperatures. Finally, K. Bonhoeffer and P. Harteck[17] obtained direct evidence for the existence of the two different forms from measurements of the heat conductivity, while A. Eucken and K. Hiller[18] obtained similar evidence from measurements of the heat capacity.

Since then a great many investigators have interested themselves in the physical properties and chemical reactions of the two forms of hydrogen. The results of this work have been summarized by A. Farkas in a monograph.[19] The following remarks are based on the discussion in Chapter II of this treatise.[20]

As shown in the previous sections the rotational energy states of the hydrogen molecule are given by the relation

$$E_r = \frac{h^2}{8\pi^2 I} K(K+1),$$

where K designates the rotational quantum number.

The corresponding eigenfunctions for a rigid rotator, as deduced in Chapter VI, are given by the relation

$$\phi_K = \epsilon^{\pm iK\eta}.$$

Like the electron, the proton also possesses a moment of spin which has the value $\pm\frac{1}{2}$ in units of $h/2\pi$, but unlike the electrons, the two protons in a hydrogen molecule can be either parallel or antiparallel in direction

[13] The values of B_v for $v = 13$ and $v = 14$ are taken from the paper by H. Beutler, *loc. cit.*

[14] *Z. Physik*, **38**, 411 (1926); **41**, 239 (1927).

[15] *Z. Physik*, **42**, 93 (1927).

[16] *Proc. Roy. Soc.* (*London*), **A115**, 483 (1927).

[17] *Naturwissenschaften*, **17**, 182 (1929); *Z. physik. Chem.*, **B4**, 113 (1929).

[18] *Z. physik. Chem.*, **B4**, 142 (1929); also K. Clusius and K. Hiller, *ibid.*, **B4**, 158 (1929).

[19] A. Farkas, " Orthohydrogen, Parahydrogen and Heavy Hydrogen."

[20] See also the discussion by the writer in Taylor's " Treatise on Physical Chemistry," Vol. 2, p. 1372.

of spin. This is due to the fact that a completely antisymmetrical function for the nuclei can be obtained with either direction of spin, and it is this phenomenon that accounts for the existence of two modifications of molecular hydrogen.

Let us consider the effect of interchanging the two coördinates of the two nuclei. Because the two nuclei are identical and symmetrically located with respect to the center of oscillation, such an interchange has no effect as regards vibrational motion. On the other hand, in the rotational eigenfunction $\epsilon^{iK\eta}$ is altered by a rotation through 180° to $\epsilon^{iK(\eta\pm\pi)}$. If K is *even*, this leaves the function unaltered, but if K is *odd*, the function changes sign. Hence, $\epsilon^{iK\eta}$ is a *symmetric function for even values of* K and *antisymmetric for odd values of* K.

As in the case of the helium atom and the hydrogen molecule, the nuclear spins combine to form three symmetric spin functions, for which $\sum m_s = -1$, 0, and 1, and one antisymmetric function for which $\sum m_s = 0$. Let ψ_S designate the symmetric spatial function $\epsilon^{iK\eta}$ and ψ_A the corresponding antisymmetric function. Then the complete eigenfunctions for the molecule are given by the eight functions:

$$\psi_A\alpha(1)\alpha(2);\quad \frac{\psi_A\{\alpha(1)\beta(2)+\alpha(2)\beta(1)\}}{\sqrt{2}};\quad \psi_A\beta(1)\beta(2);\quad \frac{\psi_S\{\alpha(1)\beta(2)-\alpha(2)\beta(2)\}}{\sqrt{2}}; \tag{A}$$

and

$$\psi_S\alpha(1)\alpha(2);\quad \frac{\psi_S\{\alpha(1)\beta(2)+\alpha(2)\beta(1)\}}{\sqrt{2}};\quad \psi_S\beta(1)\beta(2);\quad \frac{\psi_A\{\alpha(1)\beta(2)-\alpha(2)\beta(1)\}}{\sqrt{2}}. \tag{B}$$

The group (A) consists of antisymmetrical, (B), of symmetrical functions. It is possible, on the basis of experimental evidence, to decide which of these two groups actually represents the behavior of H_2.

Now observations on the intensities of lines in the band spectrum show that the lines corresponding to transitions between *odd* values of K (such as 1→1; 3→3, etc.) are *three* times as intense as those corresponding to transitions between *even* values of K (such as 0→0; 2→2, etc.). This observation can be accounted for only by assuming that the first three functions in group A represent one modification of

H_2, designated as *orthohydrogen* (by analogy with orthohelium), and that the fourth function in group A represents a second modification, designated as *parahydrogen.* That is, in the latter the nuclear spins are antiparallel while in orthohydrogen they are parallel. This is, of course, in accordance with the Pauli requirement that the complete eigenfunction for an atomic or molecular system must be of the antisymmetrical type.

Furthermore, the orthohydrogen molecule is a three-fold degenerate state. In a magnetic field the three eigenfunctions would correspond to three slightly different energy states. Hence, the *statistical* weight of the ortho modification is three times that of the para form; and we would expect that ordinary hydrogen should consist of a mixture of three parts of orthohydrogen and one part of parahydrogen. This, however, is observed to hold only at room temperature and higher, because of the following considerations.

At any temperature T, the thermodynamical equilibrium between the two forms is governed by Boltzmann's distribution law. Let N_0 denote the total number of molecules. The number in the rotational state K is given by the relation

$$N_K = N_0 p_K \epsilon^{-\frac{E_K}{kT}},$$

where p_K denotes the statistical weight. Since all para molecules have *even* values of K, and all ortho molecules, *odd* values, the ratio of the two modifications is given by

$$\beta = \frac{[p\text{-}H_2]}{[o\text{-}H_2]} = \frac{\sum_{K=\text{even}} p_K \cdot \epsilon^{-\frac{E_K}{kT}}}{\sum_{K=\text{odd}} p_K \cdot \epsilon^{-\frac{E_K}{kT}}}.$$

Since $p_K = 2K + 1$ for even values of K,

$= 3(2K + 1)$ for odd values of K,

and $$E_K = \frac{h^2}{8\pi^2 I} K(K+1) = BK(K+1),$$

it follows that

$$\beta = \frac{1 + 5\epsilon^{-\frac{6B}{kT}} + 9\epsilon^{-\frac{20B}{kT}} + \ldots}{3(3\epsilon^{-\frac{2B}{kT}} + 7\epsilon^{-\frac{12B}{kT}} + \ldots)}.$$

Now $B = 1.19 \times 10^{-14}$ gm. cm.^{2} sec^{-2} (see section 3), and $k = 1.371 \times 10^{-16}$. Hence, $B/k = 86.7$. The values of β for different values of T, as calculated by A. Farkas,[21] are shown in the following table.

T deg. K	β	Percentage of parahydrogen
20	544.8	99.82
30	32.1	96.98
40	7.780	88.61
50	3.327	76.89
75	1.077	51.86
100	0.6262	38.51
150	0.3994	28.54
210	0.3463	25.72
273	0.3357	25.13
∞	0.3333	25.00

On the basis of quantum mechanics it has been shown that no transitions (accompanied by emission or absorption of radiation) can occur between the two states. Hence, if ordinary hydrogen is cooled down to liquid air temperature, the ratio of para to ortho is not changed except in the presence of a catalyst such as charcoal. Similarly, if parahydrogen is once prepared pure it does not change spontaneously as the temperature is raised. It is the existence of this phenomenon that was used by Dennison to interpret the anomalous observations on the heat capacity of hydrogen at low temperatures.

COLLATERAL READING

The literature on band spectra is quite extensive. The most recent treatise on this subject is that of H. SPONER, "Molekülspektren und ihre Anwendungen auf chemische Probleme," two volumes, Julius Springer, Berlin, 1936.

An excellent discussion has also been given by R. DE L. KRONIG, "The Optical Basis of the Theory of Valency," The Macmillan Co., New York, 1935.

Moreover, the reader will find the topic treated by RUARK and UREY, "Atoms, Molecules and Quanta," Chapter XII, and by PAULING and WILSON, "Introduction to Quantum Mechanics," Chapter X.

The papers by C. R. JEPPESEN and H. BEUTLER (see references in chapter) should also be consulted.

[21] *Op. cit.* p. 14.

CHAPTER XIV

VALENCE BONDS; ACTIVATION AND RESONANCE ENERGY

14.1 The Homopolar Bond. In the previous chapter we discussed two alternative methods for calculating the energy of binding of two hydrogen atoms. The HL method, involving the use of atomic orbitals, is relatively simple, but leads to only approximate results; the method of molecular orbitals is more complicated, but, as shown by the work of James and Coolidge, it is possible by the use of these functions to obtain very accurate results.

For many purposes, especially where only approximate results are desired, the relation derived by Heitler and London for the energy of the shared-electron bond may be expressed conveniently as the sum of two terms in the form

$$E_S = J + K, \tag{1}$$

where J is the Coulomb energy and K is the so-called " exchange " energy. This relation is essentially the same as equation (*12*.32), in which $E_{11}/(1 + S^2) = J$, and $E_{12}/(1 + S^2) = K$. Since $| K | > | J |$ and both are negative, E_S corresponds to an energy of attraction.

Furthermore, the Pauli principle leads to the conclusion that, in the homopolar or shared-electron bond, the two electrons have opposite directions of spin. For two atoms with electrons of *similar spin moment* the energy of interaction is given by

$$E_A = J - K, \tag{2}$$

and since E_A is positive it must correspond to an energy of repulsion. Consequently, such electrons have been designated as " antibonding " in contrast to the " binding " type exhibited by two electrons of antiparallel spins.

Although, as stated previously, the concept of exchange energy is an artificial one, resulting from the mathematical arguments, equation (1) has nevertheless proved useful in the computation of activation energies. In that case it is assumed that E_S corresponds to the observed energy (that is, 100,000 cal., approximately) and that the values of J and K are in approximately the same ratio as the terms E_{11} and

E_{12} in equation (*12*.32). Thus Eyring and his associates have assumed in some of their calculations that $J/K = 1/9$, approximately.

14.2 Interaction of Two H_2 Molecules. These considerations are of special interest in deriving an approximate value for the energy of interaction of two H_2 molecules.[1] At fairly large intermolecular distances the forces acting between the molecules are of the van der Waals type, and, as shown in Chapter VIII, there is a weak force of attraction which varies as $1/r^7$. The experimental fact that H_2 condenses to a liquid only at extremely low temperatures and high pressures shows that at small intermolecular distances there is a force of repulsion.

As Penney remarks,

> The repulsive forces come about in quite a different way (than the attractive forces). In order to form a stable molecule, two hydrogen atoms interact in such a way that there is a piling up of charge density between the nuclei. This gives a stable molecule because both electrons spend an appreciable amount of their time in the region where both nuclear fields are large. Thus, if two hydrogen molecules approach one another, neither has enough electronic charge density in the region between any two protons that come together to counter-balance the repulsive forces, and the molecules therefore separate again. A more precise argument in terms of spins and energy integrals is as follows. We have seen that in order to form a stable H_2 molecule, the spins of the two electrons must be coupled to give a resultant zero. If the spins are coupled to give a resultant unity, a state with statistical weight three, strong repulsion occurs. Consider the relative orientations of the spins of two electrons located in different H_2 molecules. On the average, the two spins will behave as if they were coupled in a singlet state ($\Sigma m_s = 0$) for one-quarter of the time, and in a triplet state ($\Sigma m_s = -1, 0, +1$) for the remaining three-quarters of the time.

Consequently, the energy of interaction $E(H_2)$ of *two hydrogen atoms in different molecules* is given by the relation

$$E(H_2) = \frac{3E_A + E_S}{4} = \frac{3J - 3K + J + K}{4}$$

$$= J - \frac{K}{2}. \tag{3}$$

It should be stated that this relation is valid only as long as the intermolecular distance is at least two or three times the internuclear distance in the molecule.

By means of the Morse curve which expresses E_S as a function of internuclear distance (see equation *13*.12) it is evidently possible, with an assumed value for the ratio J/K, to calculate $E(H_2)$ as a function of intermolecular distance.

[1] W. G. Penney, " The Quantum Theory of Valency," p. 21.

14.3 Pauling's Theory of Directed Valence Bonds. Though Heitler and London showed that, on the basis of their theory of the homopolar bond, it is possible to account for the observed valences of the elements of the first two periods,[2] little or no attention was paid in this early work to the problem of directed valence bonds. The first successful efforts to treat this topic by the methods of quantum mechanics were those published by J. C. Slater[3] and, independently, by L. Pauling in a series of papers published since 1931.[4] Further investigations on the structures of specific molecules, such as H_2O and CH_4, have also been published by J. H. Van Vleck,[5] as well as by W. Heitler, E. Hückel, and G. Rumer.[6]

On the whole, the method used by L. Pauling, especially in his first papers on this topic, is more " physical " in the sense that it relies to a large extent on a geometrical interpretation of the significance of electron eigenfunctions in the formation of bonds. Even though some of the conclusions reached by Pauling may not prove to be sufficiently well founded, his method of attacking the problem of valence bonds is certainly extremely suggestive, and it is largely for this reason that the topic has been discussed in the following sections. Slater's method is more quantitative, inasmuch as it leads to approximate methods for calculating both the direction and energy of bond formation. For this reason, however, it presents greater mathematical difficulties, and the calculation becomes extremely tedious when dealing with polyatomic molecules. While Slater, Pauling, Eyring, and other investigators have developed " short-cut " rules by which roughly quantitative results can be deduced without too much labor, it is possible, in this chapter, to touch only upon the simplest aspects of the methods of solution used by the different investigators.

In his first paper, Pauling introduces the following six postulates which are to be used as a guide in the determination of relative energies and directions of different bonds in molecule formation. The first three are merely a restatement of the HL theory:

[2] See discussion by J. H. Van Vleck and A. Sherman, *Rev. Modern Phys.*, **7**, 167 (1935), especially pp. 196–197. The abbreviation V.V.S. will be used in subsequent references. This topic has also been discussed by the author in "Treatise on Physical Chemistry," by H. S. Taylor, Vol. 2, pp. 1369–1372.

[3] J. C. Slater, *Phys. Rev.*, **37**, 481; **38**, 1109 (1931).

[4] L. Pauling, *J. Am. Chem. Soc.*, **53**, 1367 (1931).

[5] References given by V.V.S.

[6] The mathematical technic used by these investigators is quite complex, but the results obtained are not essentially different from those deduced by Slater and Pauling.

1. The electron-pair bond is formed through the interaction of an unpaired electron on each of two atoms.
2. The spins of the electrons are opposed when the bond is formed, so that they cannot contribute to the paramagnetic susceptibility of the substance.
3. Two electrons which form a shared pair cannot take part in forming additional pairs.

To these Pauling adds three more rules " which are justified by the qualitative considerations of the factors influencing bond energies."

4. The main resonance (exchange) terms for a single electron-pair bond are those involving only one eigenfunction from each atom.
5. Of two eigenfunctions with the same dependence on r, the one with the larger value in the bond direction will give rise to the stronger bond, and for a given eigenfunction the bond will tend to be formed in the direction with the largest value of the eigenfunction.
6. Of two eigenfunctions with the same dependence on θ and ϕ,[7] the one with the smaller mean value of r, that is, the one corresponding to the lower energy level for the atom, will give rise to the stronger bond.

Let us consider the application of these rules to the determination of bond directions in such molecules as H_2O and NH_3. According to the Lewis-Langmuir theory of valence these molecules are represented as H:O:H and H:$\underset{\displaystyle H}{\overset{\displaystyle \cdot\cdot}{N}}$:H, respectively. The electron in a normal hydrogen atom is in the $1s$ ($n = 1$, $l = 0$) state. The electron configurations of N and O in the normal state are $(2s)^2(2p)^3$ and $(2s)^2(2p)^4$, respectively. The $2s$ electrons are paired and therefore do not take part in bond formation (except when a change occurs in quantization of the electron eigenfunction owing to bond formation, as will be discussed subsequently). Hence the electrons which act in bond formation in the case of N and O are of the type $2p$ ($n = 2$, $l = 1$).

Now, as Pauling points out, s and p eigenfunctions with the same value of n, " do not differ very much in their mean values of r, but their dependence on θ and ϕ is widely different."

For s eigenfunctions, $\Psi_{n0}(r, \theta, \phi) = S_{n0}(r) \cdot s(\theta, \phi)$.

For p eigenfunctions, $\Psi_{n1}(r, \theta, \phi) = S_{n1}(r) \cdot p(\theta, \phi)$.

The s eigenfunction is *spherically* symmetrical, and from Table 2, Chapter VII, it is seen that the normalized function has the form

$$s(\theta, \phi) = X_{00}(\theta) \cdot Z_0(\phi) = \frac{1}{\sqrt{4\pi}}.$$

[7] The symbol ϕ will be used in this chapter for the angle η used as variable in Chapter VI and subsequent discussions.

There are three p eigenfunctions:

$$p_\sigma(\theta, \phi) = X_{10}(\theta)Z_0(\phi) = \sqrt{\frac{3}{4\pi}} \cdot \cos\theta,$$

$$p_{\pi+}(\theta, \phi) = X_{11}(\theta)Z_1(\phi) = \sqrt{\frac{3}{4\pi}} \cdot \sin\theta \cdot \frac{\epsilon^{i\phi}}{\sqrt{2}},$$

$$p_{\pi-}(\theta, \phi) = X_{1,-1}(\theta)Z_{-1}(\phi) = \sqrt{\frac{3}{4\pi}} \cdot \sin\theta \cdot \frac{\epsilon^{-i\phi}}{\sqrt{2}}.$$

Since $p_{\pi+}$ and $p_{\pi-}$ are complex, it is desirable to replace them by real functions which can be represented as functions of the rectangular coördinates x, y, and z.

Replacing $\epsilon^{\pm i\phi}$ by $(\cos\phi \pm i \sin\phi)$, we obtain the real functions:

$$\left.\begin{aligned} p_x &= \sqrt{3} \cdot \sin\theta \cdot \cos\phi \\ p_y &= \sqrt{3} \cdot \sin\theta \cdot \sin\phi \\ p_z &= \sqrt{3} \cdot \cos\theta \end{aligned}\right\}, \tag{4}$$

as compared with the eigenfunction

$$s = 1,$$

where the factor $1/\sqrt{4\pi}$ has been discarded, since we are interested only in relative magnitudes.

It will be observed that p_z is identical with p_σ and that

$$p_x^2 + p_y^2 + p_z^2 = p_{\pi+}^2 + p_{\pi-}^2 + p_\sigma^2.$$

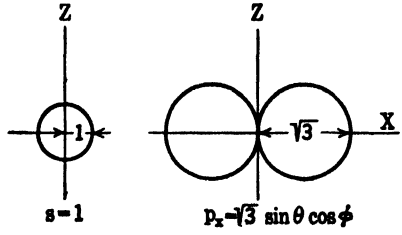


FIG. 68. s and p eigenfunctions (Pauling).

For $\phi = 0$, $p_x = \sqrt{3} \sin\theta$; and $p_y = 0$. Thus, $\sqrt{3} \sin\theta$ represents a section of the spherical function p_x in the xz-plane, in which it has its maximum value for any given value of θ. As shown in Fig. 68 this function is represented by two circles in contact at the origin and each

of diameter $\sqrt{3}$ units, as compared with the s eigenfunction which, on the same scale, is represented by a circle of unit radius.[8]

Thus $|\ p_x\ |$ consists of two spheres, with the long axis in the direction of the x-axis. The similar functions $|\ p_y\ |$ and $|\ p_z\ |$ have their long axes in the direction of the y- and z-axes, respectively. In Fig. 38, Chapter VII, the *distribution functions* are shown corresponding to these three p eigenfunctions. "From Rule 5," as Pauling observes, "we conclude that *p-electrons will form stronger bonds than s-electrons and that the bonds formed by p-electrons in an atom tend to be oriented at right angles to one another.*"

Van Vleck and Sherman[9] state the argument for this conclusion as follows:

Let us suppose there is an electron-pair bond between an s electron of some attached atom and the $p\sigma_x$ electron of the central atom. Then the exchange energy associated with this particular pair is greatest if the attached atom lies on the x-axis, since the exchange integrals will clearly be largest in absolute value if the wave functions of the two atoms overlap as much as possible. This requirement clearly demands that the attached atom be located on the axis of the dumb-bell associated with the particular electron of the central atom with which it is paired.

If a second atom is brought up, and if the pairing between p_x and the first attached atom is not broken, then clearly the only possibility is for the second atom to pair with one of the other wave functions, p_y or p_z, so that it will become located on the y- or z-axis. Hence in a molecule such as H_2O the angle between the two OH axes should be 90°. The experimental value is 106°. The departures from 90° are to be blamed upon repulsions between the attached atoms and upon sp^2 hybridization. Similarly, if the first two atoms have preempted the x and y directions a third atom tends to become located on the z-axis, so that in a molecule like NH_3 the three NH axes should make angles of 90° with each other. The NH_3 molecule is then pyramidal in structure, each axis making an angle of 54.7° with the axis of the NH_3 pyramid. The experimental value is 67°, and the discrepancy is to be attributed to the same causes as in H_2O.

14.4 Change in Quantization of Bond Eigenfunctions. In the case of a normal carbon atom, with the electron configuration $2s^2 2p^2$ (known spectroscopically as ^3P state) there are only two unpaired electrons, and this would account for the double bond in C::O. Only about 1.6 v.e. of energy (36,900 cal./mole) is required to excite one of the $2s$ electrons to a $2p$ state, and this would give three unpaired p electrons

[8] This follows from the following simple consideration:

For any point on the circle, the distance from the origin is given by $r = \sqrt{x^2 + z^2}$. Let D denote the diameter of each circle. Then it follows from the properties of any triangle inscribed in the circle on D as a base, that

$$\sin\theta = \frac{\sqrt{x^2 + z^2}}{D} = \frac{r}{D}.$$

[9] V.V.S., *op. cit.*, p. 199.

and one unpaired s electron, thus accounting for the valence of four as in CH_4. But, since the s bond is not a specially directed one, this answer cannot be sufficient. A much more satisfactory point of view is the introduction by Pauling of a new concept — that of "hybridization" of eigenfunctions to form combination functions which take the place of the single electron functions. Thus, in the case of the formation of CH_4, if the energy of interaction per H atom is greater than the difference in energy of the electron in the $2s$ and $2p$ states, then the interaction will cause the electron to be "promoted" and now we must consider each bond as being formed by the grouping together of hydrogen-like s and p eigenfunctions. As Pauling points out this criterion is satisfied for quadrivalent carbon, and he therefore proceeds to determine the zero-order eigenfunctions "which will form the strongest bonds for the case when the s-p quantization is broken." He assumes that the four bonds are each represented by one of the four combination eigenfunctions:

$$\Psi_i = a_i s + b_i p_x + c_i p_y + d_i p_z, \tag{5}$$

where $i = 1, 2, 3$, or 4, and the coefficients are subject to the orthogonality and normalization requirements[10]

$$\int \Psi_i^2 d\tau = 1 \quad \text{or} \quad a_i^2 + b_i^2 + c_i^2 + d_i^2 = 1, \tag{6}$$

and

$$\int \Psi_i \Psi_k d\tau = 0 \quad \text{or} \quad a_i a_k + b_i b_k + c_i c_k + d_i d_k = 0, \tag{7}$$

where $i \neq k$, and $i, k = 1, 2, 3, 4$.

For a single bond, we can choose the direction arbitrarily. If we take it along the x-axis, for which $p_y = p_z = 0$, the corresponding eigenfunction has the form

$$\Psi_1 = a_1 s + b_1 p_x.$$

The maximum value of this function is evidently

$$M = (\Psi_1)_{\max.} = a + b\sqrt{3}$$

(where the subscripts may be discarded).

Also since $a^2 + b^2 = 1$,

$$M = \sqrt{1 - b^2} + b\sqrt{3}.$$

[10] This point is discussed in a footnote in the paper by V. V. S., p. 202. Equation (7) is valid if the functions s, p_x, p_y, and p_z are mutually orthogonal.

Introducing the condition $dM/db = 0$, it follows that

$$b = \frac{\sqrt{3}}{2},$$

and

$$a = \tfrac{1}{2}.$$

That is,

$$\Psi_1 = \frac{1}{2}s + \frac{\sqrt{3}}{2}p_x. \tag{8}$$

This has a maximum value $M = 2$, which is considerably greater than the value 1.732 for a pure p eigenfunction. Figure 69 shows a graph of this function in the xz-plane.

A second bond function may be introduced in the same plane, of the form

$$\Psi_2 = a_2 s + b_2 p_x + d_2 p_z$$
$$= a_2 - b_2\sqrt{3}\sin\theta + d_2\sqrt{3}\cos\theta.$$

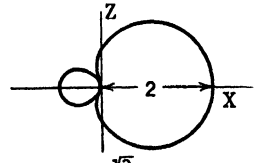


FIG. 69. Tetrahedral sp eigenfunction (Pauling).

The negative sign is due to the fact that for this bond $\cos\phi$ must be equal to -1 as compared to the value $+1$ for Ψ_1 in the xz-plane.

Applying the condition for normalization and the additional requirement

$$a_1 a_2 + b_1 b_2 = 0,$$

it is readily shown that

$$\Psi_2 = b_2\sqrt{3}(1 - \sin\theta) + \sqrt{1 - 4b_2^2}\sqrt{3}\cos\theta.$$

This will have a maximum value $M = 2$ for definite values of θ and b_2, such that $d\Psi_2/db_2 = 0$ and $d\Psi_2/d\theta = 0$. From these conditions it follows that

$$\sin\theta\cos\phi = -\tfrac{1}{3};\ \phi = 180°,\ \theta = 19°\ 28'.$$

That is, the second bond eigenfunction makes an angle of 109° 28′ with the first, "*which is just the angle between the lines drawn from the center to two corners of a regular tetrahedron.*" The actual expression for the bond function has the form

$$\Psi_2 = \frac{1}{2}s - \frac{1}{2\sqrt{3}}p_x + \frac{\sqrt{2}}{\sqrt{3}}p_z. \tag{9}$$

By similar methods it may be shown that the third and fourth eigenfunctions are

$$\Psi_3 = \frac{1}{2}s - \frac{1}{2\sqrt{3}}p_x + \frac{1}{\sqrt{2}}p_y - \frac{1}{\sqrt{6}}p_z, \tag{10}$$

and

$$\Psi_4 = \frac{1}{2}s - \frac{1}{2\sqrt{3}}p_x - \frac{1}{\sqrt{2}}p_y - \frac{1}{\sqrt{6}}p_z, \tag{11}$$

which are directed toward the other two corners of the tetrahedron.

Pauling also points out that an equivalent set of four tetrahedral eigenfunctions is

$$\begin{aligned}
\Psi_{111} &= \tfrac{1}{2}(s + p_x + p_y + p_z),\\
\Psi_{1\bar{1}\bar{1}} &= \tfrac{1}{2}(s + p_x - p_y - p_z),\\
\Psi_{\bar{1}1\bar{1}} &= \tfrac{1}{2}(s - p_x + p_y - p_z),\\
\Psi_{\bar{1}\bar{1}1} &= \tfrac{1}{2}(s - p_x - p_y + p_z),
\end{aligned}$$

which differs from the previous set by a rotation of the atom as a whole.

"This calculation," as Pauling remarks, "provides the quantum mechanical justification of the chemist's tetrahedral carbon atom, present in diamond and all aliphatic carbon compounds," as well as for a number of other tetrahedral atoms and ions. Furthermore, since "each of these tetrahedral bond eigenfunctions is cylindrically symmetrical about its bond direction, the bond energy is independent of orientation about this direction, so that there will be *free rotation about a bond.*" On the other hand, there can be *no free rotation about a double bond.*

The reader will be well repaid by studying further Pauling's original paper upon which the discussion has been based, as well as his subsequent papers. There is a wealth of material there which is of great significance for the quantum mechanical interpretation of directed valence bonds.

14.5 Slater's Treatment of Polyatomic Molecules.[11] The problem which Slater has attacked is that of calculating the electronic energy states of molecules from a consideration of the interaction of the atomic orbitals. The calculation is essentially an extension to three or more atoms of the method of secular equations discussed in section 12.5. Starting with one-electron wave functions which involve coördinates both of position and spin, zero-order wave functions are built up which are antisymmetric in the electrons (in accordance with Pauli's principle).

[11] J. C. Slater, *Phys. Rev.*, **38**, 1109 (1931).

Let $\Psi_1, \Psi_2 \ldots \Psi_i$ denote such combinations. If H is the Hamiltonian operator of the problem, then the secular equation to be solved for the energy is of the form

$$\begin{vmatrix} H_{11} - Ed_{11}, & H_{12} - Ed_{12}, \ldots & H_{1i} - Ed_{1i} \\ H_{21} - Ed_{21}, & H_{22} - Ed_{22}, \ldots & H_{2i} - Ed_{2i} \\ \hline H_{i1} - Ed_{i1}, & H_{i2} - Ed_{i2}, \ldots & H_{ii} - Ed_{ii} \end{vmatrix} = 0, \tag{12}$$

where

$$\left.\begin{aligned} H_{ij} &= \int \Psi_i H \Psi_j d\tau \\ d_{ij} &= \int \Psi_i \Psi_j d\tau \end{aligned}\right\}, \tag{13}$$

and $E_1, E_2 \ldots E_3$ represent the roots of the equation.

The degree of the secular equation will not exceed N^2, where N is the total number of electrons, although it is usually possible to separate it into a number of equations of lower degree.

Instead of attempting to review Slater's rather lengthy paper in terms of general statements, it will be more instructive to consider in detail the method used in the solution of one particular problem. Moreover, the problem we shall discuss is of importance in connection with the treatment by H. Eyring and M. Polanyi of the topic of activation energy.

Let us consider the interaction of three atoms, each containing a single electron in the s state. We can assign a one-electron wave function to each atom and we shall denote these by A, B, and C. Each of these is a function of three coördinates of position, and one of spin. Let the coördinates (of position and spin) of the first electron be denoted by 1, of the second by 2, etc. An approximate function which might be a solution for the unperturbed state of the system is the product of the individual one-electron functions of the form $A(1)\ B(2)\ C(3)$ or $A(2)\ B(1)\ C(3)$. However, there is one combination of all the different possible permutations of the functions which is the only function that is antisymmetric in the electrons. This has the form

$$\Psi = \frac{1}{\sqrt{3!}} \begin{vmatrix} A(1) & A(2) & A(3) \\ B(1) & B(2) & B(3) \\ C(1) & C(2) & C(3) \end{vmatrix}. \tag{14}$$

If there were no electron spin, this would be the only antisymmetric eigenfunction, and there would be only one corresponding eigenvalue. However, in the present case we must take into account the fact that

each electron may have either positive or negative spin. Hence, there will actually be $2^3 = 8$ different possible eigenfunctions and 8 corresponding eigenvalues. These states are indicated in the following table by Roman numerals, where a designates the part of the function A depending on the coördinates of position, and similarly b and c, while the two spin functions are indicated by + and −. The last column gives the total spin.

	Spin of a	*Spin of b*	*Spin of c*	*Total spin*
I	+	+	+	$\frac{3}{2}$
II	+	+	−	$\frac{1}{2}$
III	+	−	+	$\frac{1}{2}$
IV	−	+	+	$\frac{1}{2}$
V	−	−	+	$-\frac{1}{2}$
VI	−	+	−	$-\frac{1}{2}$
VII	+	−	−	$-\frac{1}{2}$
VIII	−	−	−	$-\frac{3}{2}$

" States I and VIII give," as Slater points out, " two of the four states of the quartet. II, III, IV give a cubic, one of whose roots gives another state of the quartet, and the other two roots give the two doublets. Similarly, V, VI, VII yield the fourth state of the quartet and the other two states of the doublet."

The antisymmetric eigenfunction corresponding to state I is evidently of the form

$$\Psi_{\mathrm{I}} = \frac{1}{\sqrt{3!}} \begin{vmatrix} a_1\alpha_1 & a_2\alpha_2 & a_3\alpha_3 \\ b_1\alpha_1 & b_2\alpha_2 & b_3\alpha_3 \\ c_1\alpha_1 & c_2\alpha_2 & c_3\alpha_3 \end{vmatrix}, \tag{15}$$

where a_1 designates the part of the function A_1 depending on the coördinates and α_1 the part which involves the spin. We shall use α and β to correspond to + and − directions of spin, respectively.

Similarly, the functions for states II and V are given by

$$\Psi_{\mathrm{II}} = \frac{1}{\sqrt{3!}} \begin{vmatrix} a_1\alpha_1 & a_2\alpha_2 & a_3\alpha_3 \\ b_1\alpha_1 & b_2\alpha_2 & b_3\alpha_3 \\ c_1\beta_1 & c_2\beta_2 & c_3\beta_3 \end{vmatrix}, \tag{16}$$

and

$$\Psi_{\mathrm{V}} = \frac{1}{\sqrt{3!}} \begin{vmatrix} a_1\beta_1 & a_2\beta_2 & a_3\beta_3 \\ b_1\beta_1 & b_2\beta_2 & b_3\beta_3 \\ c_1\alpha_1 & c_2\alpha_2 & c_3\alpha_3 \end{vmatrix}. \tag{17}$$

The eigenfunctions for the other states can be written down in a similar manner.

We now have to consider the form which each matrix element will assume in the secular equation for calculating the energy states. The energy operator for a molecule with three fixed nuclei and three electrons has the form

$$H = -\frac{h^2}{8\pi^2\mu}\nabla_i^2 - \sum\frac{Z_\alpha e^2}{r_{i\alpha}} + \sum_{\alpha\beta}\frac{Z_\alpha Z_\beta e^2}{r_{\alpha\beta}} + \sum_{ij}\frac{e^2}{r_{ij}}, \tag{18}$$

where the summations over i and j refer to the electrons, and those over α and β refer to the nuclei. The Z's represent the nuclear charges, and the r's the distances of separation.

We have in addition the fact that each function a, b, or c satisfies a one-electron S. equation

$$-\frac{h^2}{8\pi^2\mu}\nabla^2 a = (E_a - V_a)a,$$

where V_a is the potential function for the one-electron problem and E_a the energy value. If we substitute this into the operator H, we obtain the relation

$$H(a_1b_2c_3) = \left\{\sum_a (E_a - V_a) - \sum_{i\alpha}\frac{Z_\alpha e^2}{r_{i\alpha}} + \sum_{\alpha\beta}\frac{Z_\alpha Z_\beta e^2}{r_{\alpha\beta}} + \sum_{ij}\frac{e^2}{r_{ij}}\right\}(a_1b_2c_3). \tag{19}$$

The matrix elements which must be formed from the functions $\Psi_{\mathrm{I}}, \Psi_{\mathrm{II}}, \ldots \Psi_{\mathrm{VIII}}$ are of two types:

(1) those of the form $H_{\mathrm{I,\,I}}$ or $d_{\mathrm{I,\,I}}$,

and

(2) those of the form $H_{\mathrm{I,\,II}}$ or $d_{\mathrm{I,\,II}}$,

where

$$H_{\mathrm{I,\,I}} = \int \Psi_{\mathrm{I}} H \Psi_{\mathrm{I}} (d\tau d\omega)^3,$$

$$d_{\mathrm{I,\,I}} = \int \Psi_{\mathrm{I}} \Psi_{\mathrm{I}} (d\tau d\omega)^3,$$

and similarly for $H_{\mathrm{I,\,II}}$ and $d_{\mathrm{I,\,II}}$. The element of volume in coördinate space is designated by $d\tau$, and the element in spin space by $d\omega$.

Let us consider first the integral of the second type

$$H_{\mathrm{I,\,II}} = \int \Psi_{\mathrm{I}} H \Psi_{\mathrm{II}} (d\tau d\omega)^3$$

$$= \frac{1}{3!} \int \{\Sigma \pm P(a_1 b_2 c_3 \alpha_1 \alpha_2 \alpha_3)\} H \{\Sigma \pm P(a_1 b_2 c_3 \alpha_1 \alpha_2 \beta_3)\} (d\tau d\omega)^3,$$

where the symbol P denotes a permutation of the functions a, b, c over the three electrons, and the $\pm$ sign is used, depending on whether the permutation is even or odd. Since H is symmetric in the electrons, it follows that each term of the first summation yields the same result, and since there are just $3! = 6$ terms, we can write

$$H_{\mathrm{I,\,II}} = \int (a_1 b_2 c_3 \alpha_1 \alpha_2 \alpha_3) H \{\Sigma \pm P(a_1 b_2 c_3 \alpha_1 \alpha_2 \beta_3)\} (d\tau d\omega)^3.$$

There are six terms in this summation. One of the terms, *taken at random*, has the form

$$I = \int (a_1 b_2 c_3 \alpha_1 \alpha_2 \alpha_3) H (b_1 c_2 a_3 \alpha_1 \beta_2 \alpha_3)\ (d\tau d\omega)^3.$$

If we assume that there is *no interaction between the spin part of the wave function and the coördinate part*, then we can write this in the form

$$I = \int (a_1 b_2 c_3) H (b_1 c_2 a_3)\ (d\tau)^3 \int \alpha_1^2 \alpha_3^2 (d\omega)^2 \int \alpha_2 \beta_2 d\omega.$$

Now $$\int \alpha_i^2 d\omega = \int \beta_i^2 d\omega = 1 \text{ where } i = 1, 2, \text{ or } 3,$$

while $$\int \alpha_i \beta_i d\omega = 0.$$

Hence $I = 0$, and for the same reason, every term in $H_{\mathrm{I,\,II}}$ must be equal to zero. Also it follows that $d_{\mathrm{I,\,II}} = 0$. This conclusion may be stated in the more general form, that *matrix elements involving eigenfunctions belonging to different values of the total spin are equal to zero.*

We shall now introduce a convenient notation used by Slater. Instead of using the subscripts 1, 2, or 3, he indicates these by the *order* in which the functions a, b, and c are written. Furthermore, the integral sign and element of volume are omitted, so that we have

$$\int (a_1 b_2 c_3) H (b_1 c_2 a_3)\ (d\tau)^3 \equiv (abc/H/bca),$$

and

$$\int (a_1 b_2 c_3)\ (b_1 c_2 a_3)\ (d\tau)^3 \equiv (abc/1/bca).$$

The corresponding integral over both $d\tau$ and $d\omega$ will evidently vanish identically unless a, b, and c have the same spin.

We can now proceed to calculate the different matrix elements in the secular equation for the problem under discussion. Let us consider the element

$$H_{\mathrm{I,\,I}} = \int (abc) H \{\Sigma \pm P(abc)\} (d\tau)^3.$$

There are six terms, of which three are positive and three negative, as is readily seen from an inspection of the expression in equation (15) for Ψ_{I}. These terms are as follows:

$$\begin{array}{ll} abc & -cba \\ bca & -acb \\ cab & -bac \end{array}$$

Therefore,

$$H_{\mathrm{I,\,I}} = (abc/H/abc) + (abc/H/bca) + (abc/H/cab) - (abc/H/cba) - (abc/H/acb) - (abc/H/bac), \qquad (20)$$

and

$$d_{\mathrm{I,\,I}} = (abc/1/abc) + (abc/1/bca) + (abc/1/cab) - (abc/1/cba) - (abc/1/acb) - (abc/1/bac). \qquad (21)$$

Since $H_{\mathrm{I,\,II}} = H_{\mathrm{I,\,III}}$, etc. $= 0$, and similarly for $d_{\mathrm{I,\,II}}$, etc., it follows that the first row in the secular determinant is

$$H_{\mathrm{I,\,I}} - d_{\mathrm{I,\,I}} E_1 = 0,$$

while the eighth row is

$$H_{\mathrm{VIII,\,VIII}} - d_{\mathrm{VIII,\,VIII}} E_1 = 0,$$

where $H_{\mathrm{I,\,I}} = H_{\mathrm{VIII,\,VIII}}$, and $d_{\mathrm{I,\,I}} = d_{\mathrm{VIII,\,VIII}}$. Consequently, E_1 gives the energy of two of the four states of the " quartet."

That is,

$$E_1 = \frac{H_{\mathrm{I,\,I}}}{d_{\mathrm{I,\,I}}}. \qquad (22)$$

It is of interest to consider the significance of the various terms in $H_{\mathrm{I,\,I}}$ and $d_{\mathrm{I,\,I}}$.

Evidently,

$$(abc/1/abc) = 1.$$

Also $(abc/H/abc)$ is a *Coulomb integral*, since [according to equation (19)] it represents the *sum of Coulomb interactions* between all pairs of electrons in the three atoms.

A term such as $(abc/H/acb)$ in which two of the functions are interchanged represents an *exchange* integral (in this case between b and c) such as Heitler and London found in the interaction of two hydrogen atoms. If the functions on the two sides of the operator H differ by a cyclic permutation of three electrons, as in $(abc/H/cab)$ the resulting integral is zero, if the functions are orthogonal. But even if the functions are not quite orthogonal, integrals of this type have much smaller numerical values than the simple exchange integrals and may therefore be neglected in a first approximation.

Thus, if we consider the case in which one atom, say c, is at a distance from the other two atoms, so that the only permutations that count are those between a and b,

$$E_1 = \frac{(ab/H/ab) - (ab/H/ba)}{1 - (ab/1/ba)},$$

which is the energy corresponding to the antisymmetrical eigenfunction η_β in the Heitler-London theory. [See equation (*12*.35).]

Passing to the consideration of the other matrix elements, we obtain the result

$$H_{\text{II, II}} = (abc/H/abc) - (abc/H/bac). \tag{23}$$

The absence of other terms is due to the fact that we will have non-vanishing terms only in those cases in which a and b (functions with similar spin) are interchanged. All terms in which a and c or b and c are interchanged will vanish. By similar arguments it may be shown that

$$\left.\begin{aligned} H_{\text{III, III}} &= (abc/H/abc) - (abc/H/cba) \\ H_{\text{IV, IV}} &= (abc/H/abc) - (abc/H/acb) \\ H_{\text{II, III}} &= (abc/H/cab) - (abc/H/acb) \\ H_{\text{III, IV}} &= (abc/H/cab) - (abc/H/bac) \\ H_{\text{II, IV}} &= (abc/H/cab) - (abc/H/cba) \end{aligned}\right\}. \tag{24}$$

As Slater points out, " The matrices of unity are just the same with 1 substituted for H. In the last formula, we have used the fact that $(abc/H/bca) = (abc/H/cab)$. This follows from the following two steps: $(abc/H/bca) = (cab/H/abc)$, since from definition we can make any permutation of the first set of indices, if only we make an identical

permutation of the second set at the same time; and $(cab/H/abc) = (abc/H/cab)$, *since the matrices are Hermitian and real.*" [12]

The resulting cubic equation is of the form

$$\begin{vmatrix} H_{II,\,II} - d_{II,\,II}E, & H_{II,\,III} - d_{II,\,III}E, & H_{II,\,IV} - d_{II,\,IV}E \\ H_{III,\,II} - d_{III,\,II}E, & H_{III,\,III} - d_{III,\,III}E, & H_{III,\,IV} - d_{III,\,IV}E \\ H_{IV,\,II} - d_{IV,\,II}E, & H_{IV,\,III} - d_{IV,\,III}E, & H_{IV,\,IV} - d_{IV,\,IV}E \end{vmatrix} = 0. \quad (25)$$

Instead of attempting to solve this equation, Slater adopts a method which has been used successfully by other investigators subsequently. Such linear combinations of the functions Ψ_{II}, Ψ_{III}, and Ψ_{IV} are chosen as will transform the determinant into the form

$$\begin{vmatrix} H_{DD} - d_{DD}E & 0 & 0 \\ 0 & H_{AA} - d_{AA}E & H_{AB} - d_{AB}E \\ 0 & H_{AB} - d_{AB}E & H_{BB} - d_{BB}E \end{vmatrix} = 0. \quad (26)$$

In particular case the new functions A, B, and D are defined as follows

$$\left.\begin{aligned} A &= \frac{1}{\sqrt{2}}(\Psi_{II} - \Psi_{III}) \\ B &= \frac{1}{\sqrt{2}}(\Psi_{III} - \Psi_{IV}) \\ D &= \frac{1}{\sqrt{3}}(\Psi_{II} + \Psi_{III} + \Psi_{IV}) \end{aligned}\right\}. \quad (27)$$

Slater also introduces a fourth function

$$C = \frac{1}{\sqrt{2}}(\Psi_{IV} - \Psi_{II}). \quad (28)$$

It is evident that these four functions are not linearly independent, since $A + B + C = 0$. The advantage inherent in the transformation arises from the fact that each of the functions A, B, and C repre-

[12] The italics have been introduced by the writer. The reader will find it easier, in the beginning, to verify the relations given for the different matrix elements by direct expansion of the products of two determinants, such as those for Ψ_{II} and Ψ_{III}. The main point is that, wherever a product such as $\alpha_i\beta_i$ occurs, the corresponding term vanishes. This accounts for the fact that only two terms occur in each of the matrix elements, in the group of equations (24). It will be recognized that these terms are of either the Coulomb or exchange type.

sents a bond eigenfunction. Thus A corresponds to the interaction of atoms b and c, B to that of a and b, and C to that of c and a. The total spin in each of these cases is $\frac{1}{2}$, while in that of D, the total spin is $\frac{3}{2}$. Therefore, if the determinant is set up in terms of D and any two of the functions A, B, and C, it must have the form indicated in (26), since, as deduced already, matrix elements between functions corresponding to different values of the spin vanish. The result is that the determinant may be factored into a quadratic and a linear equation. The latter gives the energy of the other term of the quartet, $E = H_{DD}/d_{DD}$, while the quadratic yields the energy values of the two doublet levels.

Since $A + B + C = 0$, we obtain the result

$$\begin{aligned} H_{AA} + H_{AB} + H_{AC} &= 0, \\ H_{AB} + H_{BB} + H_{BC} &= 0, \\ H_{AC} + H_{BC} + H_{CC} &= 0. \end{aligned}$$

By eliminating from these equations, we derive the relation

$$H_{AB} = \tfrac{1}{2}(H_{CC} - H_{AA} - H_{BB}),$$

which enables us to express the non-diagonal terms in (26) in terms of the diagonal elements. The latter are readily calculated. Thus,

$$\begin{aligned} H_{AA} &= \tfrac{1}{2}(\Psi_{II} - \Psi_{III})H(\Psi_{II} - \Psi_{III}) \\ &= \tfrac{1}{2}(H_{II,\,II} + H_{III,\,III} - 2H_{II,\,III}) \\ &= (abc/H/abc) + (abc/H/acb) - (abc/H/cab) - \\ &\qquad \tfrac{1}{2}\{(abc/H/bac) + (abc/H/cba)\}. \end{aligned}$$

In a similar manner we can write down the expressions for H_{BB} and H_{CC} and, consequently, for H_{AB}. If we assume that the functions II, III, IV are approximately orthogonal, then we have the relations $d_{AA} = d_{BB} = d_{CC} = 1$, and the resulting expression for the energy of the doublets is given by

$$\begin{aligned} E = {} & (abc/H/abc) - (abc/H/cab) \pm [(abc/H/acb)^2 + \\ & (abc/H/bac)^2 + (abc/H/cba)^2 - (abc/H/acb)(abc/H/bac) - \\ & (abc/H/acb)(abc/H/cba) - (abc/H/bac)(abc/H/cba)]^{\frac{1}{2}}. \end{aligned} \tag{29}$$

In order to analyze the physical significance of this result, let us consider the three univalent atoms a, b, and c (each having a single s electron), arranged at the corners of a triangle. Let r_{ab}, etc., designate the distances between each pair of atoms; and let e_{ab}, etc., designate the total energies of binding of the corresponding pair of atoms. Using

equations (1) and (2), we can express these energies in the form

$$e_{ab} = C_{ab} \pm \gamma,$$
$$e_{bc} = C_{bc} \pm \alpha,$$
$$e_{ca} = C_{ca} \pm \beta,$$

where C_{ab} correspond to the Coulomb energy and γ to the exchange energy, with similar interpretations for the other symbols.

Each term, such as C_{ab} or γ, is a function of the internuclear distance r_{ab}, as long as the atom c is infinitely removed. We may express the energy in the form

$$e_{ab} = (ab/H/ab) \pm (ab/H/ba).$$

Without changing the values of the terms, we may also write this in the form

$$e^c_{ab} = (abc/H/abc) \pm (abc/H/bac),$$

where the presence of atom c at a considerable distance from a and b is indicated formally.

In a similar manner we may express the binding energies e^a_{bc} and e^b_{ca}. But now let us consider what happens to the value of e^c_{ab} when c is nearer to a and b. In that case, the energy of the system is represented by E in equation (29). In the latter, the first term

$$(abc/H/abc) = C_{ab} + C_{bc} + C_{ca}.$$

The second term in equation (29), $(abc/H/cab)$, which is due to the cyclic permutation of the three electrons, is known as a *multiple exchange* integral, and its actual value is small compared to the single exchange integrals, such as $(abc/H/bac)$. Hence, it may be neglected for most purposes.

Evidently $(abc/H/acb) = \alpha$, since it corresponds to the exchange energy for the molecule bc with a completely removed. Similarly, we have the relations

$$(abc/H/bac) = \gamma,$$

and

$$(abc/H/cba) = \beta.$$

Consequently, we can write equation (29) in the simplified form

$$E_{abc} = C_{ab} + C_{bc} + C_{ca} \pm (\alpha^2 + \beta^2 + \gamma^2 - \alpha\beta - \beta\gamma - \gamma\alpha)^{\frac{1}{2}}. \quad (30)$$

As shown in the discussion of the HL theory, both the Coulomb and exchange energy terms are negative for all ranges of values of internuclear distance of practical interest. Hence, $C_{ab} + \gamma$ corresponds to attraction, and since $|\gamma| > |C_{ab}|$, $C_{ab} - \gamma$ corresponds to repulsion.

For a similar reason, the positive sign in equation (30) must be used to indicate attraction, and the negative sign for repulsion.

14.6 Method of Localized Pairs. The procedure described in the previous section is obviously quite tedious. For calculating the shapes of molecules and activation energies Pauling and Van Vleck have developed a modification of Slater's method which is more convenient to use. Although results obtained in this manner are only very rough approximations, nevertheless they are of sufficient significance for the purpose of physical interpretation.

This method has been designated as that of *localized pairs* by W. G. Penney, and as that of *valence-bond wave functions* by Pauling and Wilson. The method is essentially similar to that described in a previous section in connection with Pauling's treatment of directed valence bonds. Van Vleck has also dealt along the same lines with the problem of the structure of CH_4 and H_2O.[13]

As an illustration of the application of the method of localized pairs, let us consider the interaction of an atom C with the molecule AB, the problem which has been discussed in the previous section. We shall assume that the atoms A and B are bound by a shared-electron bond (two s electrons of opposite spins) and that C contains a single s electron in the valence shell.

Let us now bring up atom C to the molecule AB in a line at right angles to the axis of the molecule. Then each atom in the molecule interacts with the atom C, and the magnitude of this interaction energy is given, according to equation (3), by the relation

$$E_{ca} + E_{cb} = C_{ca} - \frac{\beta}{2} + C_{cb} - \frac{\alpha}{2}.$$

Hence, the total energy of the system (in the lowest state) is given by

$$E^c_{ab} = C_{ab} + \gamma + C_{ca} - \frac{\beta}{2} + C_{cb} - \frac{\alpha}{2}$$

$$= \sum C + \gamma - \frac{(\alpha + \beta)}{2}. \tag{31}$$

In a similar manner we obtain, for the interaction of the atom A with the molecule BC, the total energy

$$E^a_{bc} = \sum C + \alpha - \left(\frac{\beta + \gamma}{2}\right), \tag{32}$$

[13] For details see W. G. Penney, *op. cit.*, Chapter IV, and J. H. Van Vleck and A. Sherman, *loc. cit.*

and in the case of atom B and molecule AC,

$$E_{ac}^{b} = \sum C + \beta - \frac{(\alpha + \gamma)}{2}. \tag{33}$$

These three relations obviously represent the *limiting cases* for the energy of a system of three atoms arranged at the corners of a triangle, with the three interatomic distances approximately the same. We are interested in deducing from equations (31), (32), and (33) the energy of this latter system.

Now in equation (31) the terms α and β are small compared to γ, and therefore we can write this relation in the form

$$E_{ab}^{c} = \sum C + \gamma - \frac{\alpha + \beta}{2} = \sum C + \sqrt{\gamma^2 - \alpha\gamma - \beta\gamma}, \tag{34}$$

where the right-hand side is derived from the expression in the center by omitting the terms $(\alpha^2 + \beta^2)/4$ and $\alpha\beta/2$, under the radical sign.

For E_{bc}^{a} and E_{ac}^{b} similar relations can be deduced, and since E_{abc} must involve the exchange energy terms symmetrically, we obtain the relation

$$E_{abc} = \sum C - [\tfrac{1}{2}\{(\alpha - \beta)^2 + (\beta - \gamma)^2 + (\gamma - \alpha)^2\}]^{\frac{1}{2}}, \tag{35}$$

which is identical with equation (30), except for the sign.[14] The $\pm$ sign in the latter arose from taking into account the spin degeneracy. The negative sign is used in equation (35) to indicate that we must use the negative values of the square root, since this will yield the lowest (most stable) energy state.

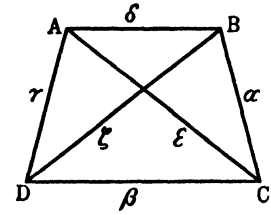


FIG. 70. Interaction energies of four atoms.

In a similar manner we may deduce the potential energy of a system of four univalent atoms with only spin degeneracy. In Fig. 70 let A, B, C, D denote the four atoms, arranged so that the interatomic distances are approximately the same. From these four atoms it is evidently possible to have the following three pairs of diatomic molecules:

$$\begin{array}{lll} AB & AD & AC \\ CD & BC & BD. \end{array}$$

Let α, β, γ, δ, ϵ, and ζ designate the exchange energies for these molecules as indicated in Fig. 70. As in the discussion of the three-atom

[14] More rigorous methods for deriving equation (35) are given by W. G. Penney, *op. cit.*, p. 77, and by Van Vleck and Sherman, *loc. cit.*

problem, we obtain the relation

$$E_{ab}^{cd} = \sum C + (\beta + \delta) - \left(\frac{\alpha + \gamma}{2}\right) - \left(\frac{\epsilon + \zeta}{2}\right),$$

and similar relations for E_{bc}^{ad} and E_{bd}^{ac}, in which the three expressions $(\alpha + \gamma)$, $(\beta + \delta)$ and $(\epsilon + \zeta)$ occur symmetrically.

By analogy with equations (31) and (35) it follows that the energy for the lowest state of the four-atom system is given by the relation

$$E = \sum C + [\tfrac{1}{2}\{(\alpha + \gamma - \beta - \delta)^2 + (\beta + \delta - \epsilon - \zeta)^2 + (\epsilon + \zeta - \alpha - \gamma)^2\}]^{\frac{1}{2}}. \tag{36}$$

14.7 Activation Energy. As first pointed out by F. London,[15] equations (35) and (36) lead to a method for calculating the activation energy of chemical reactions from a knowledge of the potential energy (Morse) curves for the molecules involved in the reactions. This point of view was adopted by H. Eyring and M. Polanyi[16] in the consideration of some simple gas reactions and has since then been extended by the former and his collaborators to the interpretation of a large number of chemical reactions.[17]

The activation energy (E_A) of a chemical reaction is determined from the temperature variation of the velocity constant k by the Arrhenius equation

$$\frac{d \ln k}{dT} = \frac{E_A}{RT^2}. \tag{37}$$

This equation may be written in the integrated form

$$k = Z\epsilon^{-\frac{E_A}{RT}}, \tag{38}$$

where Z is interpreted, on a kinetic theory basis, as the number of collisions per unit time per unit volume between reacting molecules. The exponential term corresponds, in accordance with the Boltzmann theorem, to the fraction of colliding molecules which have an energy in excess of E_A, and equation (38) is interpreted from this point of view

[15] Sommerfeld Festschrift, S. Hirzel, Leipzig, 1928, p. 104.

[16] Naturwissenschaften, **18**, 914 (1930); *Z. physik. Chem.*, **B12**, 279 (1931).

[17] Some of the more important papers are the following: H. Eyring, *Chem. Rev.*, **10**, 103 (1932); *J. Am. Chem. Soc.*, **53**, 2537 (1931); **54**, 3191 (1932); A. Sherman and H. Eyring, *J. Am. Chem. Soc.*, **54**, 2661 (1932); H. Eyring, *J. Chem. Phys.*, **3**, 107 (1935); H. Eyring, H. Gershinowitz, and C. E. Sun, *J. Chem. Phys.*, **3**, 786 (1935); H. Eyring, *Chem. Rev.*, **17**, 65 (1935); A. Wheeler, B. Topley, and H. Eyring, *J. Chem. Phys.*, **4**, 178 (1936). See also Van Vleck and Sherman, *loc. cit.* and W. G. Penney, "The Quantum Theory of Valence," Chapter V.

as indicating that only those molecules which have an energy in excess of the critical value E_A actually react.

Now let us consider the reaction

$$C + AB \rightarrow CA + B,$$

where A, B, and C are univalent atoms with s electrons. Let a, b, and c designate the corresponding single-electron functions, and let α, β, and γ designate the "exchange" energy terms for the molecules BC, CA, and AB, respectively. These symbols thus have the same significance as in equations (30) and (35).

In order that the reaction may occur, the atom C must be brought up to the molecule AB. We can do this in at least two ways, and in the first of these we let C approach AB along the surface for which $\alpha = \beta$. (For the case in which A and B are identical this will be the plane which is perpendicular to the axis of the molecule AB.) Then,

$$\begin{aligned} E^c_{ab} &= C_{ab} + C_{ca} + C_{cb} + (2\alpha^2 + \gamma^2 - \alpha^2 - 2\alpha\gamma)^{\frac{1}{2}} \\ &= \sum C + (\gamma - \alpha) \\ &= \sum C + \gamma - \frac{(\alpha + \beta)}{2}. \end{aligned} \tag{39}$$

That is, the atoms A and B will repel the atom C with an energy $-(\alpha + \beta)/2$, and at the same time the energy of binding of the system is decreased (E is made more positive).

Since both α and β increase in magnitude with decrease in the distances r_{AC} and r_{BC}, it is possible for $(\alpha + \beta)/2$ to become equal to γ, and then a reaction may occur with formation of either CA or CB, depending upon the relative magnitudes of α and β. In this case, therefore, the atom C has to acquire sufficient energy to pass over the potential energy "barrier" of magnitude $|\gamma|$, and this will be the approximate magnitude of the activation energy for the reaction. If the reaction results in the formation of the molecule CA, the binding energy for the latter is given by $C_{ca} + \beta$, and, neglecting the difference between the relatively small terms C_{ab} and C_{ca}, the total change in energy for the reaction is given by $|\gamma| - |\beta|$.

Now it is evident that the reaction must actually proceed in such a manner that the activation energy is a minimum. As shown by F. London, this condition is obtained if the atom C is brought up to the molecule AB in a direction which is the extension of the molecular axis. Assuming that the reaction leads to formation of the molecule CA, we must permit C to approach along the axis of AB in such a way

that C is closer to A than to B. In that case we may neglect the interaction energy α between B and C, and consequently,

$$E^c_{ab} = C_{ca} + C_{ba} + \sqrt{\gamma^2 + \beta^2 - \beta\gamma}.$$

The expression under the radical has a maximum value for $\beta = \gamma/2$, and under these conditions

$$E^c_{ab} = \sum C + \frac{\gamma\sqrt{3}}{2}. \tag{40}$$

That is, the atom C *decreases* the binding forces between the atoms A and B, and the energy of the system changes from the initial value $C_{ab} + \gamma$ (for C at infinity) through a maximum value $C_{ab} + C_{ac} + \gamma\sqrt{3}/2$ to a final value $C_{ac} + \beta$ (for B at infinity). The activation energy is evidently given to a first approximation by the magnitude $|\gamma| - \sqrt{3}/2\,|\gamma| = 0.134\,|\gamma|$, which is thus considerably smaller than the activation energy for the atom C brought up to AB in a direction at right angles to that of the molecular axis.

These considerations may be illustrated by means of Figs. 71[18] and 72. The energy of the system, as a function of the interatomic distances r_{CA} and r_{AB}, is represented by a *surface* in a three-dimensional coördinate system. In two dimensions, this surface may be represented by *contours* which are lines of equal energy, and in Fig. 71 the point I represents the *initial* state of system in which $(r_{AB})^0$ is the interatomic distance when C is at infinity. As C approaches AB along the direction of the line AB, the energy of the system increases, and the curves marked 1, 2, 3, etc., indicate such a series of contours. It will be observed that for large values of r_{CA} these contours are approximately parabolic in shape, since they correspond to Morse curves such as that shown in Fig. 62 for H_2. As r_{AB} becomes greater, the energy becomes more and more positive and the contours become merged in a " plateau," as indicated in the figure.

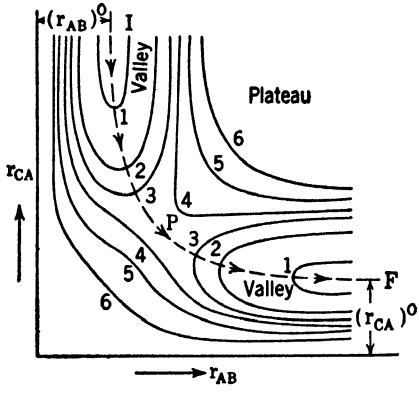


FIG. 71. Energy surface (showing contour lines) for three atoms arranged linearly.

[18] Figures similar to Fig. 71 for different reactions are shown in the paper by Eyring and Polanyi (*loc. cit.*) and in subsequent papers by Eyring and his associates. (See footnote 17.)

The *final* state of the system is represented by the point F, in which $(r_{CA})^0$ is the interatomic distance in the molecule CA, with the atom B removed to infinity. The transition from the state I to state F occurs with minimum increase in energy along the path indicated by the dashed line IPF, which is the bottom of a " valley " in the three-dimensional representation. This valley has a maximum height above I in the region P, which thus forms a " pass " or " saddle." If the system has sufficient energy to carry it from I to P, then it may either revert to state I or proceed to state F. The activation energy thus corresponds to the absolute height of the point P with respect to I.

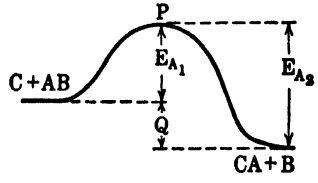


FIG. 72. Relation between energies of activation and of reaction for reversible reaction.

The change in energy of the system in passing from state I to state F is indicated graphically in Fig. 72, where E_{A_1} is the activation energy for the reaction

$$C + AB \rightarrow CA + B,$$

E_{A_2} is the activation energy for the reverse reaction, and

$$Q = E_{A_1} - E_{A_2} \tag{41}$$

is the heat of reaction.

In order to calculate the contour lines it is necessary to determine the Coulomb and exchange terms as functions of r_{CA} and r_{AB}. In his calculations Eyring has assumed that these may be derived from the potential energy curves, obtained for the diatomic molecules from band spectra or other sources, on the basis that the exchange energy term constitutes in every case by far the largest fraction of the binding energy. For instance, in the case of H_2, as mentioned in the first section of this chapter, the Coulomb energy is less than 10 per cent of the total energy of binding. That is, referring to equation (1),

$$K = 0.9E,$$

where K is the exchange energy, and E is the binding energy.

Now, as mentioned in Chapter XIII, E can be represented for a diatomic molecule, as a function of the internuclear distance, by a Morse relation of the form

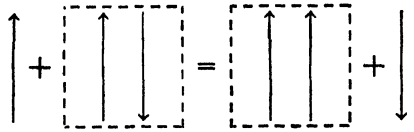
$$E(r) = -2De^{-\beta r} + De^{-2\beta r}, \tag{13.12}$$

where β and D are obtained (in general) from data on the band spectrum. Hence, K [that is, the terms α, β, γ in equation (35)] may be determined as a function of r_{AB} or r_{AC} in Fig. 71, and from such data

it is then possible to deduce a value for the minimum energy of activation for a given reaction.

A comparatively simple reaction which illustrates these considerations is that between atomic hydrogen and para- or orthohydrogen in accordance with the equation

$$H + p\text{-}H_2 = o\text{-}H_2 + H,$$



in which the directions of nuclear spin in the two molecules are indicated by the arrows. As mentioned in the previous chapter, the proportions of the two modifications in an equilibrium mixture vary with the temperature. Transition from one form to the other cannot occur spontaneously, but only through the intermediary of a third substance, as for instance atomic hydrogen.

From a study of the reaction velocity as a function of the temperature the activation energy for each of the opposing reactions may be deduced and compared with that derived from potential energy curves. Since the complete details of the experimental observations and calculations are given by A. Farkas,[19] it is only necessary in this connection to mention that the agreement obtained between the experimental and theoretical values of E_A is quite satisfactory.

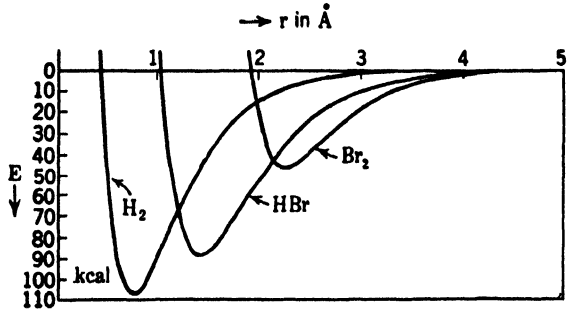


FIG. 73. Potential energy (Morse) curves for H_2, HBr and Br_2.

The reactions, involving halogens, of the type indicated by the equations

$$H + HBr \rightarrow H_2 + Br$$
$$H + Br_2 \rightarrow HBr + Br,$$

[19] "Orthohydrogen, Parahydrogen and Heavy Hydrogen," Chapter IV.

have been investigated in a similar manner by Eyring and his associates.[20] Figure 73[21] shows the Morse potential energy curves for the molecules involved in these reactions from which values of exchange terms may be derived as functions of internuclear distance. Although the agreement between the observed and calculated values of activation energy is not as good as might appear desirable, there is no doubt that this interpretation of activation energy is fundamentally sound.[22]

14.8 Resonance Energy. The method of " localized pairs " has also been applied by L. Pauling and other investigators to a type of problem in regard to molecular structures, for which chemists have been unable to obtain a definite solution by other methods.

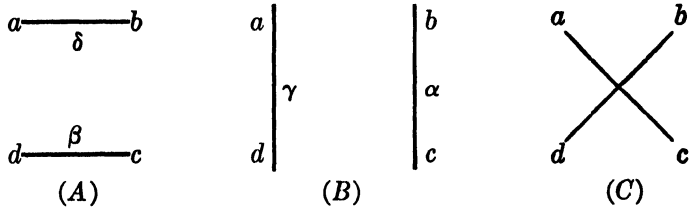
Let us consider, for instance, the case of four *equivalent univalent* atoms with only spin degeneracy. There are available $2^4 = 16$ unperturbed functions,[23] but only six of these functions have a value zero for the total spin. These will correspond to the bonds drawn thus:

Structure (A) $a:b$; $c:d$

Structure (B) $a:d$; $b:c$

Structure (C) $a:c$; $b:d$.

These may also be represented thus:



As shown by G. Rumer,[24] only two of these structures, *viz.*, (A) and (B), are independent, since (C) can be formed out of the other two. The relation between the three structures can be illustrated best by a

[20] H. Eyring, *J. Am. Chem. Soc.*, **53**, 2537 (1931).

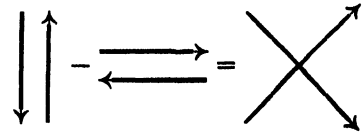
[21] Eyring and Polanyi, *loc. cit.*

[22] Although the criticism of Eyring's particular assumptions by A. S. Coolidge and H. M. James [*J. Chem. Phys.*, **2**, 811 (1934)] may be justifiable, it seems to the writer that no arguments have been advanced which would contradict the fundamental aspects of this interpretation of activation energy. The reader should also consult, in this connection, the more recent publications by Eyring and his associates on the " activated complex."

[23] J. C. Slater, *Phys. Rev.*, **38**, 1109 (1931).

[24] *Göttinger Nachr.*, p. 377, 1932; see Pauling and Wilson, " Introduction to Quantum Mechanics," p. 375.

vector diagram in the following manner:[25]



By this procedure, as Pauling points out, "Any structure can be resolved into structures involving no intersecting bonds." In consequence of this relation Pauling designates (A) and (B) as a *canonical set*, and describes the system by the molecular orbital

$$\Psi = a\psi_A + \sqrt{1 - a^2} \cdot \psi_B. \tag{42}$$

The interpretation of this relation is identical with that assigned to analogous relations in Chapter XII. The system behaves as if constituted of two types of structures, and the relative probabilities of occurrence of the structures are given by the ratio $a^2/(1 - a^2)$.

Let α, β, γ, δ denote the exchange terms for the molecules, as indicated above. If we neglect the exchange terms between diagonal atoms and let $\delta = \beta = \gamma = \alpha$, then equation (36) leads to a binding energy for the system

$$E = \sum C + 2\alpha. \tag{43}$$

Comparing this with the value $C_{ab} + \alpha$ which would be valid for either of the structures (A) or (B) above, it is seen that the very possibility of representing the system by a wave function of the form given in equation (42) leads to an *increase in the energy of binding*. This increase, which is given by α in this case, is known as the *resonance energy*. Resonance occurs whenever a molecule can be represented by more than one equivalent structural formula, that is, when there is more than one way of joining up the bonds. The resonance energy is a measure of the *increase in stability* of the system which results from the fact that the molecule can apparently "resonate" between two or more equally reasonable structures. The analogy between this phenomenon and the increase in energy of two oscillators in virtue of their interaction is evident.

For systems involving more than four equivalent atoms, special methods for the approximate determination of the resonance energy have been developed by L. Pauling[26] and also by H. Eyring and G. E. Kimball.[27]

[25] L. Pauling, *J. Chem. Phys.*, **1**, 280 (1933).

[26] *Loc. cit.*, also "Introduction to Quantum Mechanics," pp. 374–382.

[27] *J. Chem. Phys.*, **1**, 239 (1933).

One of the most interesting problems to which these methods have been applied is that of the structure of benzene, a problem which was first treated by E. Hückel.[28] Chemists have argued for a long period regarding the proper representation of the molecule, and, as is well known, different structures have been suggested which are shown in Fig. 74*a*. Of these, the best known is that of Kekulé, but very good reasons have also been advanced for the adoption of the other structural formulas.

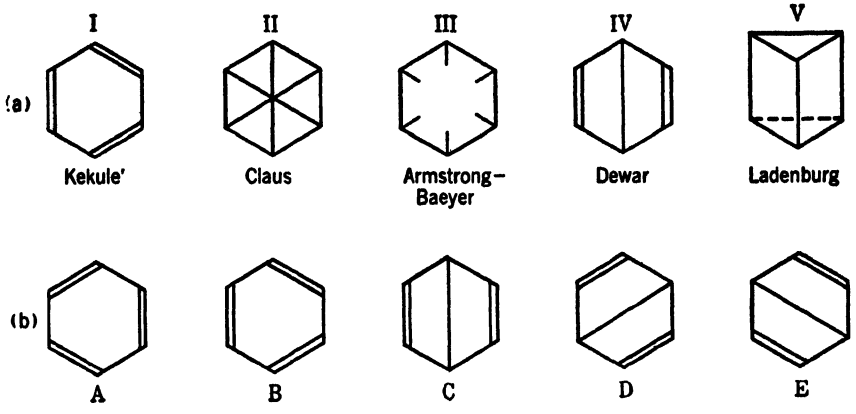


FIG. 74. (*a*) Suggested valence-bond structures for benzene ring.
(*b*) Five canonical valence-bond structures for benzene, leading to resonance phenomenon.

As shown by Pauling the whole argument is resolved by the assumption that in the normal state the molecule may be represented by a mixed eigenfunction of the form

$$\Psi = a(\psi_A + \psi_B) + b(\psi_C + \psi_D + \psi_E), \tag{44}$$

where the different component functions refer to the five canonical structures shown in Fig. 74*b*. Equation (44) indicates, first, that the molecule may be regarded as possessing the properties of each of the five structures; second, that structures A and B are equally probable; third, that C, D, and E are also equally probable; and last, that the group $A + B$ has a different probability from that of the group $C + D + E$. Thus, we may consider that the fraction of the time during which the molecule exists in state A is $a^2/(2a^2 + 3b^2)$, and the rest of the time in the other states. In other words, the solution given by quantum mechanics to a problem which has vexed chemists for many decades is that everyone who suggested a solution was neither completely right nor completely wrong.

[28] *Z. Physik*, **70**, 204 (1931).

According to L. Pauling and G. W. Wheland,[29] the values of the coefficients in equation (44) are $a = 1$, $b = 0.4341$. Designating the single exchange integral for two *adjacent* hydrogen atoms in the ring by α, the conclusions deduced for the value of the resonance energy are shown in the following table.

	Total Energy	*Resonance Energy*
Single Kekulé structure	$Q + 1.5\alpha$	0.0
Resonance between A and B	$Q + 2.4\alpha$	0.9α
Resonance among all five structures	$Q + 2.6055\alpha$	1.1055α

The exchange energy term for a single Kekulé structure would be 1.5α, where α is about 1.5 v.e. (= 34,580 cal.), and the added resonance energy term is 1.1055α, that is, about 38,200 cal./mole, " of which about 80 per cent is due to the Kekulé structures alone."[30]

In a more recent publication, E. Hückel[31] has reached the conclusion that in equation (44) the values of the coefficients are $a = 0.450$, $b = -0.147$. On the other hand, he deduces the same value for the resonance energy as Pauling and Wheland.

The latter have also applied the same method to the calculation of resonance energies for other ring compounds, such as naphthalene and anthracene, and a large number of other organic molecules. Although the absolute values derived for the resonance energy may be doubtful, owing to the rather approximate methods of calculation, the resonance phenomenon is unquestionably of fundamental importance in leading to increased stability of molecules.

[29] *J. Chem. Phys.*, **1**, 362 (1933).

[30] From their very accurate determinations of the heats of organic reactions, G. B. Kistiakowsky and his co-workers, *J. Am. Chem. Soc.*, **58**, 146 (1936), deduce a value of about 36,000 cal. for the resonance energy of benzene. However, they find that in general the resonance calculations of Pauling are not in satisfactory agreement with the experimentally observed values. This, of course, merely indicates that it is at present impossible by the methods used in quantum mechanics to derive even approximately correct values of resonance energies.

[31] *J. Phys. Radium*, **6**, 347 (1935).

CHAPTER XV

QUANTUM MECHANICS THEORY OF RADIATION

15.1 Classical Theory of Radiation. The electromagnetic theory of Clerk Maxwell led to the conclusion that radiant energy is propagated as a wave motion. The simplest model of a source for the production of such waves is an electric dipole with fixed axis and periodically variable moment M, where

$$M = ex = ex_0 \cos 2\pi\nu t. \tag{1}$$

In this equation x_0 represents the maximum amplitude of the oscillating charge, and ν, the frequency of oscillation.

Such a dipole is known as an *electrical oscillator* or *Hertzian vibrator* and finds its macroscopic representation in a short linear antenna such as is used for the radiation of very high-frequency radio waves.

According to electromagnetic theory an electric charge moving with uniform velocity does not radiate energy. If, however, the charge is accelerated, the instantaneous rate at which energy is emitted is given by the relation

$$-\frac{dE}{dt} = \frac{2e^2}{3c^3}\left(\frac{d^2x}{dt^2}\right)^2, \tag{2}$$

where x is the coördinate along which motion occurs (the axis of the dipole), and c is the velocity of light.

From equations (1) and (2) it follows that for the linear harmonic oscillator

$$-\frac{dE}{dt} = \frac{2(2\pi\nu)^4}{3c^3} e^2x_0^2 \cos^2(2\pi\nu t).$$

The average rate of emission of energy per single oscillator is

$$J = -\left(\overline{\frac{dE}{dt}}\right) = \frac{e^2x_0^2}{3c^3}(2\pi\nu)^4,$$

since the average value of $\cos^2(2\pi\nu t)$ is $\frac{1}{2}$, and it will be noted that the frequency of the radiation emitted is *identical* with that of the oscillator.

In general, the charge will not be concentrated at any one point, but will be represented by a distribution function $\rho(x, y, z)$. For the case

of a negative charge distribution about a positive charge located at the origin (as in atomic systems), the electric moment of the charge distribution will be represented by the *vector quantity* (see Appendix IV)

$$\mathbf{M} = e\int \mathbf{r}\rho(x, y, z)dxdydz \cdot \cos(2\pi\nu t), \tag{3}$$

where $\mathbf{r}$ is the vector distance from the origin. That is, $\mathbf{r}$ indicates both the *magnitude* of the distance of the element of volume and the *direction* in which this distance is measured. If we represent the components of the moment along the three axes of coördinates by $\mathbf{M}_x$, $\mathbf{M}_y$, and $\mathbf{M}_z$, the intensity of the radiation for which the direction of polarization is along the x-axis is given by

$$J_x = \frac{(2\pi\nu)^4}{3c^3} \cdot \mathbf{M}_x^2, \tag{4}$$

and similar relations will apply to the components for which the directions of polarization are along the other two axes of coördinates. Consequently, the average rate of emission of radiation for all three components is given by

$$J_x + J_y + J_z = \frac{(2\pi\nu)^4}{3c^3}\{\mathbf{M}_x^2 + \mathbf{M}_y^2 + \mathbf{M}_z^2\}. \tag{5}$$

In the case of a charge moving in some form of periodic orbit, as was postulated in the Bohr theory, the motion along any coördinate q can be represented by Fourier series of the form

$$\begin{aligned} q = {} & A_0 + A_1 \cos 2\pi\nu t + \ldots + A_n \cos 2\pi\nu n t + \ldots \\ & + B_1 \sin 2\pi\nu t + \ldots + B_n \sin 2\pi\nu n t + \ldots, \end{aligned} \tag{6}$$

where, if q is real, A_n and B_n are *real* magnitudes.

We can also express q in the exponential form

$$q = \sum_n C_n \epsilon^{2\pi i n\nu t}, \tag{7}$$

where n has all integral values (including 0) from $-\infty$ to $+\infty$. This is necessary, since, as shown in Chapter II,

$$\begin{aligned} & C_n\epsilon^{2\pi i n\nu t} + C_{-n}\epsilon^{-2\pi i n\nu t} \\ = {} & C_n \cos 2\pi n\nu t + iC_n \sin 2\pi n\nu t + \\ & C_{-n} \cos 2\pi n\nu t - iC_{-n} \sin 2\pi n\nu t \\ = {} & A_n \cos 2\pi n\nu t + B_n \sin 2\pi n\nu t, \end{aligned}$$

where

$$\begin{aligned} 2C_n &= A_n - iB_n \\ 2C_{-n} &= A_n + iB_n. \end{aligned}$$

Consequently, $\bar{C}_n = C_{-n}$, and

$$C_n \cdot \bar{C}_n = C_n \cdot C_{-n} = |C_n|^2 = \frac{(A_n^2 + B_n^2)}{4}.$$

From equation (6) it follows that, for this case, the intensity of the radiation, that is, the average value J of the energy emitted per unit time, is given by the relation

$$J = \frac{(2\pi\nu)^4 e^2}{3c^3} \cdot \overline{q^2}, \tag{8}$$

where $\overline{q^2}$ = average value of q^2. From the orthogonality relations for the trigonometric functions it follows that

$$\overline{q^2} = \sum_n (A_n^2 + B_n^2), \tag{9}$$

and consequently,

$$J = \frac{(2\pi\nu)^4 e^2}{3c^3} \sum_n 4C_n^2, \tag{10}$$

where n assumes all integral values from $+1$ to $+\infty$.

Again the fundamental frequency ν and the harmonics $n\nu$ will appear in the radiation, and the magnitude of the coefficient $|C_n|^2$ will be a measure of the relative intensity of the corresponding harmonic in the radiation emitted.

The Bohr theory of the origin of spectral lines represented a radical departure from this point of view since it postulated that the spectral frequencies are obtained as a result of transitions between two levels in accordance with the relation

$$\nu_{nm} = \frac{(E_n - E_m)}{h}. \tag{11}$$

Consequently, optical frequencies are not identical, on the basis of Bohr's theory, with frequencies of revolution. However, when we are dealing with orbits of large quantum number, the frequency of the radiation emitted as a result of transitions between adjacent orbits becomes approximately the same as that of the motion in either orbit.[1] That is, we have the approximate relation

$$\nu = \Delta n \cdot \omega,$$

where ν is the optical frequency, ω is the frequency of revolution of the

[1] See discussion in Taylor's Treatise, pp. 1179–1181.

charge, and $\Delta n = 1, 2$, etc. Thus Δn corresponds to the integer n in equations (9) and (10), and it is seen that under these conditions (that is, in which ν is associated with transitions between orbits of large quantum numbers) the deductions of quantum theory are in agreement with those expected on the basis of classical theory.

As emphasized by Bohr this correspondence or coincidence in the values of spectral frequencies as calculated from the two points of view cannot be accidental. The coincidence must also extend to amplitude (or intensity) of emitted radiation and direction of polarization, and Bohr showed that on this basis it is possible to deduce selection rules for the probabilities of transition between energy levels. This idea, which has been designated by Bohr as the Correspondence Principle, has also been carried over into quantum mechanics, and may be stated thus: For systems in which the product of momentum p and associated coördinate q is very large compared to h, the deductions of quantum mechanics must tend to become identical with those derived from classical theory. It will be recognized that the condition $pq \gg h$ is fulfilled by Bohr orbits of large quantum number.

15.2 The Schroedinger Equation Involving Time. In order to interpret the method developed in quantum mechanics for calculating dipole moments corresponding to transitions, it is necessary to consider once more the derivation of the S. equation as given in Chapter II.

Starting with the differential equation for a wave propagation in the form

$$\nabla^2\psi = \frac{1}{\nu^2\lambda^2}\frac{\partial^2\psi}{\partial t^2}, \tag{12}$$

where ψ, the "amplitude," is a function of the configuration space coördinates and of t, a solution was postulated of the form

$$\psi(q_i, t) = \phi(q_i)\epsilon^{-2\pi i\nu t}, \tag{13}$$

where q_i denotes the generalized coördinates, and ν designates the frequency of the de Broglie wave. This solution satisfies equation (12), and if we multiply each term in the S. equation

$$\nabla^2\phi + \frac{8\pi^2\mu}{h^2}(E - V)\phi = 0$$

by $\epsilon^{-2\pi i\nu t}$, we obtain the differential equation

$$\nabla^2\psi - \frac{8\pi^2\mu}{h^2}V\psi = -\frac{8\pi^2\mu}{h^2}E\psi. \tag{14}$$

Let us now assume the relation

$$E = h\nu, \tag{15}$$

thus defining the frequency ν, about which nothing definite was stated previously. The justification for this procedure depends, of course, upon the validity of the subsequent deductions, since we cannot determine, by any direct experimental methods, the value of ν for the de Broglie wave. Moreover, it is evident, since we always observe *differences* in energy, that we might also assume

$$E = \text{constant} + h\nu.$$

Actually de Broglie postulated the relation

$$\nu = \frac{E + \mu_0 c^2}{h},$$

where μ_0 is the so-called " rest-mass " of the particle and $E + \mu_0 c^2$ is the absolute energy on the basis of the special theory of relativity. However, it is evidently simpler to postulate equation (15).

It follows from equations (13) and (15) that

$$\frac{d\psi}{dt} = -\frac{2\pi i}{h} E\psi,$$

and, hence, equation (14) becomes

$$\nabla^2\psi - \frac{8\pi^2\mu}{h^2} \cdot V\psi = \frac{h}{2\pi i} \frac{8\pi^2\mu}{h^2} \frac{d\psi}{dt} = -\frac{4\pi\mu i}{h} \frac{d\psi}{dt}. \tag{16}$$

For the conjugate complex function $\bar{\psi}$, we have the similar relation

$$\nabla^2\bar{\psi} - \frac{8\pi^2\mu}{h^2} V\bar{\psi} = \frac{4\pi\mu i}{h} \frac{d\bar{\psi}}{dt}. \tag{17}$$

Hence,

$$\frac{d}{dt}(\bar{\psi}\psi) = \frac{h}{4\pi\mu i}(\psi\nabla^2\bar{\psi} - \bar{\psi}\nabla^2\psi). \tag{18}$$

If now we multiply both sides by $d\tau$, the element of volume in the configuration space, and integrate, we obtain the relation

$$\frac{d}{dt}\int \bar{\psi}\psi d\tau = \frac{h}{4\pi\mu i}\int (\psi\nabla^2\bar{\psi} - \bar{\psi}\nabla^2\psi)d\tau. \tag{19}$$

But in accordance with Green's theorem (see Appendix IV)

$$\int (\psi\nabla^2\bar{\psi} - \bar{\psi}\nabla^2\psi)d\tau = \int \left(\psi \frac{d\bar{\psi}}{dn} - \bar{\psi}\frac{d\psi}{dn}\right)d\sigma,$$

where the integral on the right-hand side represents integration over a " surface " bounding the " volume " in which the left-hand integration is carried out, and $d\psi/dn$ is the gradient in the direction of the normal to this surface. At the limits of the configuration space, ψ and $\bar{\psi}$ must vanish. (This is one of the conditions which the S. eigenfunctions must satisfy.) Hence, the right-hand side of equation (19) must vanish identically, and consequently,

$$\frac{d}{dt}\int \bar{\psi}\psi d\tau = 0. \tag{20}$$

That is,

$$\int \bar{\psi}\psi d\tau = \text{constant}. \tag{21}$$

It is this result that permits us to interpret $\bar{\psi}\psi$ as a density distribution function and to normalize ψ so that

$$\frac{1}{N^2}\int \bar{\psi}\psi d\tau = 1.$$

From equation (16) it follows that for $V = 0$, that is, for particles moving in a zero field

$$\nabla^2\psi = -\frac{4\pi\mu i}{h}\frac{d\psi}{dt}. \tag{22}$$

This equation has the same form as the diffusion equation [see Appendix IV, equation (41)], with $-\frac{h}{4\pi\mu i} = \frac{ih}{4\pi\mu} = D$, the " diffusion constant."

Now let us consider the solution of equations (16) or (17). Because of the linearity of the equation, the *general* solution has the form

$$\psi = \sum_n \psi_n = \sum_1^\infty c_n\phi_n\epsilon^{-\frac{2\pi i E_n t}{h}}, \tag{23}$$

where ϕ_n and E_n are the eigenfunction and eigenvalue associated with the nth state, and c_n is an arbitrary *constant.*

Similarly,

$$\bar{\psi} = \sum_n \bar{c}_n\bar{\phi}_n\epsilon^{\frac{2\pi i E_n t}{h}}. \tag{24}$$

That is, ψ (or $\bar{\psi}$) is represented by a *superposition* of wave functions, each of which corresponds to a definite energy state.

Hence,

$$\bar{\psi}\psi = \sum_n \bar{c}_n c_n \bar{\phi}_n \phi_n + \sum_{n,m} \bar{c}_m c_n \bar{\phi}_m \phi_n \epsilon^{-\frac{2\pi i(E_n - E_m)t}{h}}, \tag{25}$$

where the second summation extends over all values of n and m which are *not identical.* If we multiply each term in this equation by $d\tau$, and integrate over the configuration space, the integrals involving the exponential terms vanish, since

$$\int \bar{\phi}_n \phi_m d\tau = 1 \text{ for } n = m,$$
$$= 0 \text{ for } n \neq m.$$

It follows from equation (25) that

$$\int \bar{\psi}\psi d\tau = \sum |\, c_n \,|^2. \tag{26}$$

We can interpret the quantities $|\, c_n \,|^2$ as the *relative probabilities* of finding the particle in the corresponding alternative states, and as in equation (21) we can normalize the function ψ by the condition

$$\sum |\, c_n \,|^2 = 1.$$

Let us now consider the exponential terms in equation (25). The individual terms can be arranged in pairs of the form

$$\bar{c}_m c_n \bar{\phi}_m \phi_n \epsilon^{-2\pi i \nu_{nm} t} + c_m \bar{c}_n \phi_m \bar{\phi}_n \epsilon^{-2\pi i \nu_{mn} t}, \tag{27}$$

where
$$h\nu_{nm} = E_n - E_m = -h\nu_{mn}. \tag{28}$$

Since $\bar{c}_m c_n \bar{\phi}_m \phi_n$ and $c_m \bar{c}_n \phi_m \bar{\phi}_n$ are complex conjugates, we can write the expression in (27) in the form

$$\rho(\cos 2\pi\nu_{nm} t + i \sin 2\pi\nu_{nm} t), \tag{29}$$

where ρ is a charge distribution function.

It is thus evident that $\bar{\psi}\psi$ is a periodic function in t, with a series of "interference" frequencies given by equation (28). These are, however, the very frequencies which occur as transition frequencies in the Bohr theory. Hence, we conclude that $\bar{\psi}\psi$ *is a quantity which may function as a source of electromagnetic vibrations,* giving rise to the monochromatic radiations postulated in the Bohr theory.

On the basis of this deduction it becomes possible to translate the classical theory of radiation, as outlined in the previous section, in terms of wave mechanics.

15.3 Quantum Mechanics Expressions for Dipole Moment and Intensity of Radiation. Referring to the classical expression for $\mathbf{M}_x$, that is, the x-component of the dipole moment, given in equation (1), we shall write the quantum mechanics analog for the emission of frequency $\nu_{nm} = (E_n - E_m)/h$, in the form[2]

$$(\mathbf{M}_{nm})_x = e\int \bar{\psi}_n x \psi_m d\tau + e\int \bar{\psi}_m x \psi_n d\tau$$
$$= e\int x\bar{\phi}_n\phi_m \epsilon^{-2\pi i\nu_{nm}t} d\tau + e\int x\bar{\phi}_m\phi_n \epsilon^{2\pi i\nu_{nm}t} d\tau. \qquad (30)$$

It will be recognized that the two integrals on the right-hand side of the last equation constitute a pair of complex conjugated matrix elements which may be designated by the symbols x_{nm} and x_{mn} respectively. Hence,

$$\frac{d^2(\mathbf{M}_{nm})_x}{dt^2} = -e(2\pi\nu_{nm})^2(x_{nm} + x_{mn}). \qquad (31)$$

In order to derive the corresponding expression for $(J_{mn})_x$, the intensity of radiation emitted by the x-component of the dipole moment, it is necessary to calculate the time-average value of the square of the expression on the right-hand side of equation (31).

Evidently,

$$|\,(x_{nm} + x_{mn})\,|^2 = 2\int x\bar{\phi}_n\phi_m d\tau \int x\bar{\phi}_m\phi_n d\tau$$
$$+ \left|\int x\bar{\phi}_n\phi_m d\tau\right|^2 \epsilon^{-4\pi i\nu_{nm}t}$$
$$+ \left|\int x\bar{\phi}_m\phi_n d\tau\right|^2 \epsilon^{4\pi i\nu_{nm}t}. \qquad (32)$$

Now the average value of $\int \epsilon^{\pm n\pi i\nu t} dt$ over one or more complete cycles is the same as that of $\int \cos n\pi\nu t dt$ or $\int \sin n\pi\nu t dt$, and is evidently equal to zero. Furthermore, as mentioned already, the matrix elements x_{mn} and x_{nm} are complex conjugates irrespective of the exponential terms. Hence, the relation for intensity of radiation assumes the form

$$(J_{nm})_x = \left[\frac{d^2(\mathbf{M}_{nm})_x}{dt^2}\right]^2 = \frac{2e^2}{3c^3}(2\pi\nu_{nm})^4 2\left|\int x\bar{\phi}_n\phi_m d\tau\right|^2. \qquad (33)$$

[2] The derivation that follows is to be regarded as a somewhat plausible argument for the use of equation (30). For a more rigorous derivation the reader should consult Pauling and Wilson, " Introduction to Quantum Mechanics," Chapter XI.

This equation is the wave mechanics analog of equation (4). From this and the similar expressions for the y- and z-components of the dipole moment, it is possible, as in equation (5), to calculate the components of the radiation intensity with respect to different directions of polarization.

While these relations enable us to calculate the magnitudes of the radiation intensities in terms of the matrix elements built up out of the corresponding eigenfunctions, they are specially useful in deducing so-called selection rules.

If in any given case the matrix element x_{mn} is found to vanish identically we conclude that the corresponding component is not present in the emitted radiation, and if this same result is derived for y_{mn} and z_{mn} then we conclude that the corresponding transition cannot occur.

The integrals x_{nm} and x_{mn} are evidently similar to integrals encountered in the discussion of the perturbation theory, of the type

$$f_{mn} = \int \bar{\psi}_m f \psi_n d\tau,$$

where f is a function of the coördinates. Thus, in Chapter IX, f is identical with V, the potential energy function. These integrals were designated matrix elements, and it is evident, since the matrix elements defined by equation (31) are complex conjugates, that the matrix is, again, of the Hermitian type.

15.4 Selection Rule for Linear Harmonic Oscillator. A relatively simple case in which the matrix elements associated with a transition from one state to another may be calculated readily is that of the *linear harmonic oscillator* which was discussed in Chapter V. It was shown there that, omitting the time factor,

$$x_{mn} = \int_{-\infty}^{\infty} x\epsilon^{-x^2} H_n(x) H_m(x) dx \tag{34}$$

is equal to zero, unless $m = n \pm 1$. For $m = n + 1$,

$$x_{n+1,n} = \sqrt{\frac{n+1}{2}}, \tag{35}$$

and for $m = n - 1$,

$$x_{n,n-1} = \sqrt{\frac{n}{2}}. \tag{36}$$

That is, in the case of the linear harmonic oscillator only those transitions can occur for which the quantum number changes by $+1$ or -1.

We have thus derived a *selection rule* for transitions between energy states of a linear harmonic oscillator. Furthermore, the radiation emitted in a transition between adjacent states must be linearly polarized and the direction of polarization must coincide with that of the motion of the oscillating particle.

According to equation (33), the intensity of the radiation associated with the transition $n \to n-1$ is given by the relation

$$J_{n,n-1} = \frac{(2\pi\nu)^4 \cdot 4e^2}{3c^3} \left| \frac{x_{n,n-1}}{\sqrt{b}} \right|^2, \tag{37}$$

where according to equation (5.7)

$$b = \frac{4\pi^2 \mu\nu}{h}.$$

Hence:

$$J_{n,n-1} = \frac{(2\pi\nu)^4}{3c^3} \cdot \frac{2ne^2}{b}. \tag{38}$$

Similarly, it follows that

$$J_{n+1,n} = \frac{(2\pi\nu)^4}{3c^3} \cdot \frac{2(n+1)e^2}{b}. \tag{39}$$

It is of interest to compare these results with the conclusions from classical theory on the basis $E_n = (n + \frac{1}{2})h\nu$. In that case, the value of the maximum amplitude for the state n is $\sqrt{2n+1} \cdot x_0$, where x_0 is the maximum amplitude for the state $n = 0$, while that for the state $\sqrt{2n-1}$ is $\sqrt{2n-1} \cdot x_0$. Therefore, the classical value for the intensity of a transition from the level n to the level $n-1$ would be between

$$(2n+1)\frac{(2\pi\nu)^4}{3c^3} \cdot e^2x_0^2 \quad \text{and} \quad (2n-1)\frac{(2\pi\nu)^4}{3c^3} \cdot e^2x_0^2.$$

For *large values* of n it is evident that this agrees with the conclusions stated in equations (38) and (39).

15.5 Selection Rules for Quantum Numbers m and l. In Chapter VI it was shown that the eigenfunction for the rigid rotator with fixed axis is given by

$$Z_m = \frac{\epsilon^{\pm im\eta}}{\sqrt{2\pi}}, \tag{40}$$

where $|\, m \,| = 0, 1, 2,$ etc.

The x-component of the center of gravity of the charge consists of terms of the type

$$x_{nm} = \int x Z_m \bar{Z}_n d\eta \cdot \epsilon^{-2\pi i(\nu_m - \nu_n)t}. \tag{41}$$

If we assume that the rotator is represented by a charge e moving on a circle of radius r, we obtain the relations

$$\left.\begin{aligned} x = r\cos\eta &= \frac{r}{2}(\epsilon^{i\eta} + \epsilon^{-i\eta}) \\ y = r\sin\eta &= \frac{r}{2i}(\epsilon^{i\eta} - \epsilon^{-i\eta}) \end{aligned}\right\}. \tag{42}$$

Omitting the time factor, it follows from equations (40), (41), and (42) that

$$x_{nm} = \frac{r}{4\pi}\left[\int \epsilon^{i(m-n+1)\eta} d\eta + \int \epsilon^{i(m-n-1)\eta} d\eta\right], \tag{43}$$

where the integration is over the range $0 \leqslant \eta \leqslant 2\pi$.

Similarly,

$$y_{nm} = \frac{r}{4\pi i}\left[\int \epsilon^{i(m-n+1)\eta} d\eta - \int \epsilon^{i(m-n-1)\eta} d\eta\right]. \tag{44}$$

In equation (43) the first integral vanishes unless

$$m - n + 1 = 0,$$

and the second integral vanishes unless

$$m - n - 1 = 0.$$

In either case the integral is equal to 2π, and

$$x_{nm} = \frac{r}{2} \text{ for } n = m \pm 1, \tag{45}$$

while

$$\left.\begin{aligned} y_{nm} &= \frac{r}{2i} \text{ for } n = m + 1 \\ &= -\frac{r}{2i} \text{ for } n = m - 1 \end{aligned}\right\}. \tag{46}$$

We have thus derived a selection rule for the rotator, that only those transitions can occur for which $\Delta m = \pm 1$. In other words, in the case

of the rigid rotator with fixed axis, the only transitions permitted are those *between adjacent states.*

To determine the direction of polarization of the emitted radiation, we note that

$$\begin{aligned}\mathbf{M}_x(m-1, m) &= e\int xZ_m\bar{Z}_{m-1}d\eta \cdot \epsilon^{2\pi i\nu t} + e\int x\bar{Z}_mZ_{m-1}d\eta \cdot \epsilon^{-2\pi i\nu t} \\ &= er\cos 2\pi\nu t.\end{aligned}$$

Similarly, the relations are deduced

$$\begin{aligned}\mathbf{M}_x(m+1, m) &= er\cos 2\pi\nu t, \\ \mathbf{M}_y(m-1, m) &= -er\sin 2\pi\nu t, \\ \text{and} \qquad \mathbf{M}_y(m+1, m) &= er\sin 2\pi\nu t.\end{aligned}$$

These relations show that the radiation emitted and absorbed is *circularly polarized,* since the x- and y-components of the electric moment are equal for each transition. For the spontaneous transition $m+1 \to m$, the light emitted is circularly polarized in the direction of the rotation, and for the transition $m-1 \to m$, the light *absorbed* is also that which is polarized in the direction of rotation.

It also follows from equation (45) that

$$J_x(m \pm 1, m) = \frac{16\pi^4\nu^4e^2r^2}{3c^3}. \tag{47}$$

We shall now consider the *rigid rotator with free axis.* The eigenfunctions for this system, as derived in Chapter VI, are given by

$$Y_{l,m}(\theta, \eta) = \left\{\frac{(2l+1)\,(l-m)!}{4\pi(l+m)!}\right\}^{\frac{1}{2}} \epsilon^{\pm im\eta}P_l^m(\cos\theta). \tag{48}$$

The dipole moment corresponding to transition from state l, m to state l', m' is given by

$$\mathbf{M}(l, m, l', m') = \int \mathbf{r}Y_{lm}\bar{Y}_{l'm'}dxdydz,$$

in which the term involving t has been omitted, and $\mathbf{r}$ is a vector which is defined in terms of x, y, and z.

Hence, the x-component of the dipole moment is given by

$$\mathbf{M}_x = \frac{er}{N^2}\int_0^\pi \sin\theta \cdot P_l^m(\cos\theta) \cdot P_{l'}^{m'}(\cos\theta) \cdot \sin\theta d\theta \cdot$$

$$\int_0^{2\pi} \cos\eta\epsilon^{im\eta}\epsilon^{-im'\eta}d\eta, \tag{49}$$

where $1/N$ designates the coefficient in equation (48). Similarly,

$$\mathbf{M}_y = \frac{er}{N^2}\int_0^\pi \sin\theta \cdot P_l^m(\cos\theta) \cdot P_{l'}^{m'}(\cos\theta) \cdot \sin\theta d\theta \cdot \int_0^{2\pi} \sin\eta \cdot \epsilon^{im\eta}\epsilon^{-im'\eta} d\eta, \tag{50}$$

and

$$\mathbf{M}_z = \frac{er}{N^2}\int_0^\pi \cos\theta \cdot P_l^m(\cos\theta) \cdot P_{l'}^{m'}(\cos\theta) \cdot \sin\theta d\theta \cdot \int_0^{2\pi} \epsilon^{im\eta}\epsilon^{-im'\eta} d\eta.$$

Let us first consider $\mathbf{M}_z$. Since the integral involving η vanishes for $m \neq m'$, we need to consider, in calculating $\mathbf{M}_z$, only those transitions for which $\Delta m = 0$, and consequently the integral involving η is equal to 2π. Hence,

$$\mathbf{M}_z = \frac{er}{N^2} \cdot 2\pi \int_0^\pi \cos\theta \cdot \sin\theta\, P_l^m(\cos\theta) P_{l'}^{m'}(\cos\theta) d\theta. \tag{51}$$

Now in any treatise on Legendre functions the following recursion formula is deduced:

$$(2l+1)\cos\theta \cdot P_l^m(\cos\theta) = (l+m)P_{l-1}^m(\cos\theta) + (l-m+1)P_{l+1}^m(\cos\theta).$$

Substituting from this relation in equation (51) it follows from the orthogonal properties of the function $P_l^m(\cos\theta)$ that $\mathbf{M}_z$ vanishes unless $l' = l \pm 1$.

By evaluating $\mathbf{M}_x$ and $\mathbf{M}_y$ it is found that light polarized along these axes is emitted (or absorbed) only when $m' = m \neq 1$. This conclusion has already been deduced in the consideration of the rigid rotator with fixed axis.

We have thus deduced the following selection rules: Only those transitions can occur for which

$$\Delta m = 0 \quad \text{or} \quad \pm 1, \quad \text{and} \quad \Delta l = \pm 1.$$

In the case of the total quantum number n, it is found that there is no similar selection principle. Transitions are permitted between any pair of values of this number.

15.6 Wave Packets. Group Velocity. In this section we shall consider some further consequences of the S. equation (18) involving time. It will be recognized that this differential equation must in some manner give us the same kind of information about the function $\bar{\psi}\psi$ as is obtained in classical mechanics from the equations of motion. As shown in Chapter IV, the latter may be expressed in terms of the Hamiltonian function $H = H(p_i, q_i)$ by the canonical equations

$$\frac{dq_i}{dt} \equiv \dot{q}_i = \frac{\partial H}{\partial p_i}, \tag{52}$$

and

$$\frac{dp_i}{dt} \equiv \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{53}$$

Given the initial values of p_i, q_i and the form of the function $H(p_i, q_i)$ it is possible, by solving these canonical equations, to predict the magnitudes of the variables p_i and q_i for any subsequent value of t. In view of the validity of the Principle of Indeterminacy it is evidently impossible, as stated previously, to carry out such precise calculations in connection with systems for which the product $p_i q_i$ is of the same order of magnitude as h. However, we do know that, as the magnitude of the product $p_i q_i$ increases to values which are considerably greater than h, the methods of classical mechanics become more and more applicable. This is, in fact, the essential significance of the Correspondence Principle. Therefore, in the formulation of quantum mechanics we must be guided by the criterion that in the limit of large quantum numbers, that is, for large values of $\oint p_i dq_i$, the results obtained by the new theory must tend to become identical with those derived by classical mechanics.

The essential difference between the classical and quantum mechanics is, of course, due to the fact that the purely corpuscular aspect of motion is no longer adequate for the description of atomic systems. The observations on reflection and refraction of electrons make it necessary to regard the behavior of these particles as somehow associated with "waves," the de Broglie waves. Furthermore, as Heisenberg demonstrated, this involves a radical revision of our notions regarding the significance of coördinates and momentum variables in the case of atomic systems, and he formulated this new point of view in his Principle of Indeterminacy.

On the basis of this principle a homogeneous beam of electrons travelling in a given direction, that of the x-coördinate, possesses a definite

momentum p, and we represent this state of affairs by the eigenfunction

$$\psi = A\epsilon^{i\alpha x}\epsilon^{-2\pi i\nu t},$$

where $$\alpha = \frac{2\pi}{\lambda} = \frac{2\pi p}{h} = \frac{2\pi\sqrt{2\mu E}}{h},$$

and $$\nu = \frac{E}{h}.$$

The real part of the complex function ψ corresponds to a sine or cosine curve extending from $-\infty$ to $+\infty$. When we inquire regarding the instantaneous position of the electron, the answer given by quantum mechanics is that, while we are completely ignorant of this coördinate, it is possible to predict the *probability* of occurrence of an electron per unit length of its path.

Similarly, we may define the instantaneous position x of a particle. But in that case, the velocity may be any value from zero to infinity. How then can we represent a " localized " particle? In the language of quantum mechanics this means that the form of the function $\bar{\psi}\psi$ is to be such that it has physically significant values over only *an extremely narrow range* of values of x. The best illustration of such a function is the probability curve represented by the relation

$$y = b\epsilon^{-c^2x^2}. \tag{54}$$

The plot of this function has the form shown in Fig. 24 for the eigenfunction of the linear harmonic oscillator in the state $n = 0$.[3] This curve is symmetrical about the origin, and for $x = \pm 1/c$, the value of y is $1/e$th, that is, 0.3678 of its value for $x = 0$. Evidently, $1/c$ is a measure of the " spread " of the curve. The greater the value of c, the smaller the value of $|x|$ for which the value of y decreases to any given fraction of its value for $x = 0$.

On the basis of Fourier's theory, any function which is continuous and single valued over a range of values of x may be represented by a series involving cosine or sine functions.[4] Each of these functions will represent a stationary wave of definite wave length λ, which corresponds, therefore, to a definite momentum p. Now if we wish to represent the probability function $\bar{\psi}\psi$ by such a series it is found that the number of terms required to represent the function ψ with any degree

[3] See Chapter V, supplementary note 1, also J. W. Mellor's " Higher Mathematics."

[4] See Chapter VI, supplementary note 1.

of approximation increases with increase in the coefficient c in equation (54). In terms of quantum mechanics this means that, the more we attempt to minimize Δx in the representation of $\psi\bar{\psi}$, the greater the range of values of p which are required in forming the " wave packet " for $\psi(x)$.

Thus, the sense in which the term "wave packet" is used in quantum mechanics is essentially the following. In order to obtain a function $\psi\bar{\psi}$ which shall localize the particle in a relatively narrow region of the configuration space, we represent ψ by a summation of trigonometric terms each of which corresponds to the amplitude (as a function of coördinates and time) of a de Broglie wave of definite wave length and frequency. That is, we can always represent a given ψ by the series

$$\psi = \sum a_n \epsilon^{i\alpha x},$$

in which the values of the coefficients are obtained by means of the relation

$$a_n = \int_{x_1}^{x_2} \psi \epsilon^{-i\alpha x} dx,$$

and the actual number of terms will depend upon how accurately it is desired to approximate the function ψ in a region $x_2 > x > x_1$.

We shall now deduce the relation between the velocity of the particle, regarded as a corpuscle, and the velocity of propagation of the de Broglie waves which constitute the wave packet. For this purpose let us consider the simple wave train

$$y = y_0 \cos 2\pi(\nu t - \sigma x), \tag{55}$$

in which ν = frequency, and $\sigma = 1/\lambda$ = wave number. If we superpose on this wave another wave of the same amplitude (y_0) but slightly different frequency $\nu + \Delta\nu$ and slightly different wave number $\sigma + \Delta\sigma$, the amplitude of the resulting motion is given by the relation

$$\begin{aligned} y &= y_0 \cos 2\pi(\nu t - \sigma x) + y_0 \cos 2\pi[(\nu + \Delta\nu)t - (\sigma + \Delta\sigma)x] \\ &= 2y_0 \cos 2\pi(\nu t - \sigma x) \cos \pi(\Delta\nu t - \Delta\sigma x). \end{aligned} \tag{56}$$

The equation indicates a train of waves of frequency $\Delta\nu$ and wave number $\Delta\sigma$ which is modulated by the other wave of much higher frequency and higher wave number. The maximum amplitude travels with a velocity v defined by the relation

$$v = \frac{x}{t} = \frac{\Delta\nu}{\Delta\sigma} = \frac{d\nu}{d\sigma} \text{ in the limit.} \tag{57}$$

On the other hand, the individual waves travel with velocity

$$u = \frac{\nu}{\sigma}.$$

The latter is designated the *phase velocity*, and v, the *group velocity*. Equation (57) is well known in the theory of wave motion and is often written in the form

$$\frac{1}{v} = \frac{d\left(\frac{\nu}{u}\right)}{d\nu} = \frac{1}{u} - \frac{\nu}{u^2} \cdot \frac{du}{d\nu},$$

which gives the relation between the two magnitudes u and v. Evidently, $v = u$ only if $du/d\nu = 0$, that is, if the phase velocity is independent of the frequency. This is true of electromagnetic waves in a vacuum. Therefore, under these conditions, the phase and group velocities are equal. On the other hand, the group velocity of waves propagated in water, is one-half the phase velocity. That is, the individual waves travel faster than the group.

In a wave motion, energy is transferred only at the speed defined by the group velocity. Consequently, in the case of de Broglie waves, the group velocity is identical with that of the particle. The phase velocity is a fictitious magnitude, since it is impossible to observe for de Broglie waves a frequency ν.

This conclusion may also be stated thus: For the corpuscle regarded as a wave motion

$$\nu = \frac{E}{h}; \quad \lambda = \frac{h}{p}; \quad \sigma = \frac{p}{h}. \tag{58}$$

But in accordance with equation (57), the group velocity is given by the relation

$$v = \frac{d\nu}{d\sigma} = \frac{dE}{dp}, \tag{59}$$

which corresponds to the canonical equation

$$\dot{q} = \frac{\partial H}{\partial p}. \tag{60}$$

If now we represent E as a function of p for any set of particles in motion, then the corpuscular velocities are determined for each value of E by the value of the slope dE/dp.

15.7 Relation between Eigenfunction Representation and Particle Velocity. We shall now consider some further deductions from the S. equation involving time. For this purpose it is convenient to write equation (16) in the form

$$\frac{\partial \psi}{\partial t} = \frac{ih}{4\pi\mu}\left(\frac{\partial^2 \psi}{\partial x^2} - \alpha V \psi\right), \tag{61}$$

and equation (17) in the form

$$\frac{\partial \bar{\psi}}{\partial t} = -\frac{ih}{4\pi\mu}\left(\frac{\partial^2 \bar{\psi}}{\partial x^2} - \alpha V \bar{\psi}\right), \tag{62}$$

where $\psi = \psi(x, t)$ and $\alpha = 8\pi^2\mu/h^2$.

Hence, we derive, as before, equation (18), that is, the relation

$$\frac{d}{dt}(\bar{\psi}\psi) = \frac{ih}{4\pi\mu}\left(\bar{\psi} \cdot \frac{\partial^2 \psi}{\partial x^2} - \psi \cdot \frac{\partial^2 \bar{\psi}}{\partial x^2}\right),$$

or

$$\frac{d\rho}{dt} = -\frac{h}{4\pi\mu i} \cdot \frac{d}{dx}\left(\bar{\psi} \cdot \frac{\partial \psi}{\partial x} - \psi \cdot \frac{\partial \bar{\psi}}{\partial x}\right), \tag{63}$$

where $\rho = \bar{\psi}\psi$, is the " density."

In order to realize the significance of this equation let us consider the motion of particles along the x-coördinate. Let ρ designate the density per unit length, and j, the " current," that is, the rate at which particles pass a given point. If $d\rho/dt = 0$, there is obviously no variation in j with distance. However, if the current flowing into the section Δx is different from that flowing out of it, there must be a change in density, in accordance with the relation

$$\frac{d\rho}{dt} = -\frac{dj}{dx}. \tag{64}$$

Comparing (64) with (63) it follows that the *current density* at x is defined in quantum mechanics by the relation

$$j = \frac{h}{4\pi\mu i}\left(\bar{\psi} \cdot \frac{\partial \psi}{\partial x} - \psi \cdot \frac{\partial \bar{\psi}}{\partial x}\right). \tag{65}$$

Hence, the average velocity in the x-direction is given by

$$v = \int_{-\infty}^{\infty} j\,dx = \frac{h}{4\pi\mu i}\int_{-\infty}^{\infty}\left(\bar{\psi} \cdot \frac{\partial \psi}{\partial x} - \psi \cdot \frac{\partial \bar{\psi}}{\partial x}\right)dx. \tag{66}$$

More generally, if ψ is a function of the configuration space,

$$j = \frac{h}{4\pi\mu i}(\bar{\psi}\,\mathrm{grad}\psi - \psi\,\mathrm{grad}\bar{\psi}), \tag{67}$$

where the " gradient " is the rate of change of ψ along the normal to the " surface " at any point. (This result may be derived directly from the divergence theorem which is discussed in Appendix IV.)

Equation (66) may also be derived from the definition of the velocity

$$v = \frac{d\bar{x}}{dt} = \frac{d}{dt}\left(\int_{-\infty}^{\infty} \bar{\psi} x \psi dx\right), \tag{68}$$

where $\bar{x}$ designates the average value of x.

Hence,

$$\frac{d\bar{x}}{dt} = \int x\left(\bar{\psi}\cdot\frac{\partial\psi}{\partial t} + \psi\cdot\frac{\partial\bar{\psi}}{\partial t}\right)dx.$$

(In these and subsequent equations, it will be understood that the limits of integration are $\pm\infty$.) Substituting from equations (61) and (62) the last relation becomes

$$\frac{d\bar{x}}{dt} = \frac{ih}{4\pi\mu}\int x\left(\bar{\psi}\cdot\frac{\partial^2\psi}{\partial x^2} - \psi\cdot\frac{\partial^2\bar{\psi}}{\partial x^2}\right)dx.$$

Now,

$$\frac{d}{dx}\left(x\bar{\psi}\cdot\frac{\partial\psi}{\partial x}\right) = x\cdot\frac{\partial\bar{\psi}}{\partial x}\cdot\frac{\partial\psi}{\partial x} + x\bar{\psi}\cdot\frac{\partial^2\psi}{\partial x^2} + \bar{\psi}\cdot\frac{\partial\psi}{\partial x}\cdot$$

Hence,

$$\int x\bar{\psi}\cdot\frac{\partial^2\psi}{\partial x^2}\cdot dx = x\bar{\psi}\cdot\frac{\partial\psi}{\partial x}\Big]_{-\infty}^{\infty} - \int\bar{\psi}\cdot\frac{\partial\psi}{\partial x} - \int x\cdot\frac{\partial\bar{\psi}}{\partial x}\cdot\frac{\partial\psi}{\partial x}\cdot dx.$$

But, $x\bar{\psi}\cdot(\partial\psi/\partial x)$ vanishes at the limits. Hence, we derive the relation[5] for the particle velocity in the form

$$v = \frac{d\bar{x}}{dt} = \frac{h}{4\pi\mu i}\int\left(\bar{\psi}\cdot\frac{\partial\psi}{\partial x} - \psi\cdot\frac{\partial\bar{\psi}}{\partial x}\right)dx,$$

which is identical with equation (66).

[5] Instead of the coefficient $h/(4\pi\mu i)$ we can evidently use the coefficient $-ih/(4\pi\mu)$.

It follows that the average value of the momentum is given by the relation

$$\bar{p} = \mu v = \frac{h}{4\pi i}\int\left(\bar{\psi}\cdot\frac{\partial\psi}{\partial x} - \psi\cdot\frac{\partial\bar{\psi}}{\partial x}\right)dx$$

$$= \frac{h}{4\pi i}\int\left\{\frac{d}{dx}(\bar{\psi}\psi) - \psi\cdot\frac{\partial\bar{\psi}}{\partial x} - \psi\cdot\frac{\partial\bar{\psi}}{\partial x}\right\}dx$$

$$= \frac{ih}{2\pi}\int\psi\cdot\frac{\partial\bar{\psi}}{\partial x}\cdot dx \qquad (69a)$$

$$= \frac{h}{2\pi i}\int\bar{\psi}\cdot\frac{\partial\psi}{\partial x}\cdot dx. \qquad (69b)$$

The last equation is one which follows, as shown in section 2.7, from the operator method for deriving the S. equation, since

$$\bar{p} = \int\bar{\psi}p\psi dx = \frac{h}{2\pi i}\int\bar{\psi}\cdot\frac{\partial\psi}{\partial x}\cdot dx.$$

From equations (69a), (61), and (62) it follows that

$$\frac{d^2\bar{x}}{dt^2} = \frac{ih}{2\pi\mu}\frac{d}{dt}\int\psi\cdot\frac{\partial\bar{\psi}}{\partial x}\cdot dx$$

$$= \frac{ih}{2\pi\mu}\int dx\left[\frac{\partial\psi}{\partial t}\cdot\frac{\partial\bar{\psi}}{\partial x} + \psi\cdot\frac{d}{dt}\frac{\partial\bar{\psi}}{\partial x}\right]$$

$$= \frac{ih}{2\pi\mu}\int dx\left[\frac{\partial\psi}{\partial t}\cdot\frac{\partial\bar{\psi}}{\partial x} + \psi\cdot\frac{d}{dx}\frac{\partial\bar{\psi}}{\partial t}\right]$$

$$= \frac{ih}{2\pi\mu}\cdot\frac{ih}{4\pi\mu}\int dx\left[\left(\frac{\partial^2\psi}{\partial x^2} - \alpha V\psi\right)\frac{\partial\bar{\psi}}{\partial x}\right.$$

$$\left. - \psi\cdot\frac{d}{dx}\left(\frac{\partial^2\bar{\psi}}{\partial x^2} - \alpha V\bar{\psi}\right)\right].$$

Now,
$$\frac{d}{dx}\left(\psi\cdot\frac{\partial^2\bar{\psi}}{\partial x^2}\right) = \psi\cdot\frac{d}{dx}\left(\frac{\partial^2\bar{\psi}}{\partial x^2}\right) + \frac{\partial\psi}{\partial x}\cdot\frac{\partial^2\bar{\psi}}{\partial x^2}.$$

Also, $\psi\cdot(\partial^2\bar{\psi}/\partial x^2)$ vanishes at the limits. Hence,

$$\frac{d^2\bar{x}}{dt^2} = -\frac{h^2\alpha}{8\pi^2\mu^2}\int\bar{\psi}\cdot\frac{\partial V}{\partial x}\cdot\psi dx.$$

That is,

$$\mu \cdot \frac{d^2\bar{x}}{dt^2} = -\left(\overline{\frac{\partial V}{\partial x}}\right), \tag{70}$$

where the dash over the right-hand bracket indicates that the average value of $\partial V/\partial x$ is to be used. Evidently, equation (70) corresponds to Newton's second law of motion, and it has thus been shown that the relations of quantum mechanics are in agreement, for extremely narrow wave packets, with those deduced in classical mechanics.

15.8 Spreading out of Wave Packet in Time. It is of interest, in order to illustrate Heisenberg's principle, to consider the inferences that may be deduced from this principle regarding the variation with time, in the form of a wave packet. Equation (25) shows that for a particle in a zero field, confined to move along the x-coördinate,

$$\frac{\partial \psi}{\partial t} = D \cdot \frac{\partial^2 \psi}{\partial x^2}, \tag{71}$$

where

$$D = \frac{ih}{4\pi\mu}. \tag{72}$$

The solution of this differential equation gives $\psi = \psi(x, t)$ and therefore makes it possible to determine the behavior of $\bar{\psi}\psi$ as t is increased from $t = 0$.

For real values of D, equation (71) is the well-known partial differential equation for diffusion, and the simplest form of solution is that given by the expression[6]

$$\psi(x, t) = \frac{C}{\sqrt{D(t+\tau)}} \epsilon^{-\frac{x^2}{4D(t+\tau)}}, \tag{73}$$

where τ and C are constants. If we designate the exponential factor by ϕ, we obtain the following relations:

$$\frac{\partial \psi}{\partial t} = \left\{ -\frac{1}{2\sqrt{D}(t+\tau)^{\frac{3}{2}}} + \frac{x^2}{4D^{\frac{3}{2}}(t+\tau)^{\frac{5}{2}}} \right\} C\phi,$$

$$\frac{\partial \psi}{\partial x} = -\frac{x \cdot C\phi}{2D^{\frac{3}{2}}(t+\tau)^{\frac{3}{2}}},$$

$$\frac{\partial^2 \psi}{\partial x^2} = \left\{ -\frac{1}{2D^{\frac{3}{2}}(t+\tau)^{\frac{3}{2}}} + \frac{x^2}{4D^{\frac{5}{2}}(t+\tau)^{\frac{5}{2}}} \right\} C\phi,$$

[6] Byerly, " Fourier's Series," pp. 93–95.

and it is seen that the solution in (73) satisfies the differential equation (71).

For the case in which D is *real*, it is found that this solution corresponds to the diffusion of C particles which originally ($t = -\tau$) were concentrated in the plane $x = 0$.

In the present case, let us define $2D\tau$ as a *real* positive magnitude by the relation $2D\tau = a^2$. Then we can write equation (73) in the form

$$\psi = \frac{C}{\sqrt{a^2 + \frac{iht}{2\pi\mu}}} \epsilon^{-\frac{x^2}{2(a^2 + iht/2\pi\mu)}}.$$

Also,

$$\bar{\psi} = \frac{C}{\sqrt{a^2 - \frac{iht}{2\pi\mu}}} \epsilon^{-\frac{x^2}{2(a^2 - iht/2\pi\mu)}}.$$

Hence,

$$\bar{\psi}\psi = \frac{C^2}{a\sqrt{a^2 + \left(\frac{ht}{2\pi\mu a}\right)^2}} \epsilon^{-\frac{x^2}{a^2 + (ht/2\pi\mu a)^2}}. \tag{74}$$

Evidently, $\bar{\psi}\psi$ represents the same type of Gauss error function as that defined in equation (54), in which $1/c^2$ is replaced by

$$a_t^2 = a^2 + \left(\frac{ht}{2\pi\mu a}\right)^2. \tag{75}$$

For $t = 0$, the absolute value of $\psi\bar{\psi}$ for $x = 0$ is given by C^2/a^2. With increase in the absolute value of t, a_t increases, with the result that the maximum value of $\bar{\psi}\psi$ decreases and the packet spreads out more and more. This may also be described as follows:

The probability of occurrence of the particle in the range $-x_0 \leqslant x \leqslant x_0$ is given by the integral

$$P = \int_{-x_0}^{x_0} \bar{\psi}\psi dx.$$

Equation (74) shows that, with increase in $|t|$, the value of P for the *same* limits decreases. Furthermore, since $\bar{\psi}\psi$ is symmetrical with respect to $t = 0$, it follows that this increasing uncertainty in the value of P *extends to both the past and future.* That is, no matter how exactly the value of P may be defined for a narrow range at $t = 0$, the gradual

spreading out of the packet with time will result in a correspondingly increasing uncertainty in the value of P.

While this argument has been applied in the previous paragraphs to the case in which the average value of x remains zero, a similar argument applies to the case of motion in a field of force. For an instructive illustration of such a calculation the reader should consult the discussion by C. G. Darwin[7] in which he has demonstrated that the law of motion of a packet in a gravitational field, as deduced by the methods of quantum mechanics, is identical with that deduced in ordinary mechanics.

COLLATERAL READING

1. The Correspondence Principle. For further discussion consult Ruark and Urey, "Atoms, Molecules and Quanta," Chapter VI; Slater and Frank, "Theoretical Physics," Chapter XXX.

2. Wave Mechanics Theory of Radiation. Condon and Morse, Chapter III; Ruark and Urey, pp. 542–47; J. Frenkel, "Wave Mechanics," Vol. 1, p. 124, et seq.

3. Intensities of Spectral Lines. A very comprehensive discussion is given by H. Bethe, "Handbuch der Physik," XXIV/1, p. 429 et seq.

4. Wave Packets and Group Velocity. The presentation in the chapter is based on the discussion by J. Frenkel, "Einführung in die Wellenmechanik," Julius Springer, Berlin, 1929, pp. 60–65. The reader should also consult Gurney, "Elementary Quantum Mechanics," pp. 72–75, and A. Landé, "Vorlesungen über Wellenmechanik," Academische Verlagsgesellschaft m.b.H., Leipzig, 1930.

[7] *Proc. Roy. Soc.*, **A117**, 258 (1928). See also E. H. Kennard, *J. Franklin Institute*, **207**, 47 (1929).

APPENDIX I

TREATISES AND REVIEW PUBLICATIONS ON QUANTUM MECHANICS

Special References

1. BETHE, H., Quantenmechanik der Ein- und Zwei-Elektronenprobleme, Handbuch der Physik, Vol. XXIV, Part I, pp. 273–560, Julius Springer, Berlin, 1933.
2. CONDON, E. U., and MORSE, P. M., "Quantum Mechanics," McGraw-Hill Book Co., New York, 1929.
3. CONDON, E. U., and SHORTLEY, G. H., "The Theory of Atomic Spectra," Macmillan Co., New York, 1935.
4. DARROW, K. K., "Introduction to Wave Mechanics," *Bell System Tech. J.*, **6**, 653–701 (1927). (Also German translation.)
5. DIRAC, P. A. M., "The Principles of Quantum Mechanics," Clarendon Press, Oxford, second edition, 1935.
6. FRENKEL, J., "Wave Mechanics," Vol. I, Elementary Theory (1932); Vol. II, Advanced General Theory (1934), Clarendon Press, Oxford.
7. GURNEY, R. W., "Elementary Quantum Mechanics," Cambridge University Press, Cambridge, 1934.
8. HAAS, A., "Wave Mechanics and the New Quantum Theory," translated by L. W. CODD, D. Van Nostrand Co., New York, 1928.
9. HEISENBERG, W., "The Physical Principles of the Quantum Theory," Chicago University Press, Chicago, 1930.
10. HUND, F., Allgemeine Quantenmechanik des Atom- und Molekülbaues, "Handbuch der Physik," Vol. XXIV, Part I, pp. 561–694, Julius Springer, Berlin, 1933.
11. KEMBLE, E. C., "General Principles of Quantum Mechanics, Part I," *Phys. Rev. Suppl.*, Vol. I, p. 157 (1929); E. C. KEMBLE and E. L. HILL, "Part II," *Rev. Modern Phys.*, **2**, 1 (1930).
12. LAUE, M. v., Korpuskular- und Wellentheorie, "Handbuch der Radiologie," Vol. 6, Part I, pp. 1–114, Akademische Verlagsgesellschaft, m.b.H., Leipzig, 1933.
13. MOTT, N. F., "An Outline of Wave Mechanics," Cambridge University Press, Cambridge, 1930.
14. PAULING, L., and WILSON, E. B., Jr., "Introduction to Quantum Mechanics," McGraw-Hill Book Co., New York, 1935.
15. RUARK, A. E., and UREY, R. C., "Atoms, Molecules and Quanta," McGraw-Hill Book Co., New York, 1930.
16. SOMMERFELD, A., "Wave Mechanics," translated by H. L. BROSE, E. P. Dutton & Co., New York, 1930.

General References

1. Joos, G., "Theoretical Physics," translated by I. M. Freeman, G. E. Stechert & Co., New York, 1934.
2. Lindsay, R. B., and Margenau, H., "Foundations of Physics," John Wiley & Sons, New York, 1936.
3. Pauling, L., and Goudsmit, S. A., "The Structure of Line Spectra," McGraw-Hill Book Co., New York, 1930.
4. Slater, J. C., and Frank, N. H., "Introduction to Theoretical Physics," McGraw-Hill Book Co., New York, 1933.
5. White, H. E., "Introduction to Atomic Spectra," McGraw-Hill Book Co., New York, 1934.

APPENDIX II

VALUES OF PHYSICAL CONSTANTS[1]

Molar volume of ideal gas V_M.

At 0°C. and 760 mm. Hg ($= 1.0132 \times 10^6$ dyne/cm.^{2}), $V_M = 22{,}414$ cm.^{3}

Molar gas constant,

$$\begin{aligned} R &= 8.3136 \times 10^7 \text{ erg} \cdot \text{deg}^{-1} \cdot \text{mole}^{-1} \\ &= 8.2046 \times 10^{-2} \text{ liter} \cdot \text{atmos} \cdot \text{deg}^{-1} \cdot \text{mole}^{-1} \\ &= 1.9864 \text{ cal}_{15} \cdot \text{deg}^{-1} \cdot \text{mole}^{-1}. \end{aligned}$$

Faraday constant, $F = 96{,}490$ abs-coul $\cdot$ g-equiv.$^{-1}$

Velocity of light, $c = 2.99796 \times 10^{10}$ cm $\cdot$ sec.$^{-1}$

Electronic charge, $e = 4.770 \times 10^{-10}$ abs-es-units.

$e/c = 1.591 \times 10^{-20}$ abs-em-units.

Specific electronic charge, $e/m_0 = 1.761 \times 10^7$ abs-em-unit $\cdot$ g^{-1}

Mass of electron, $m_0 = 9.035 \times 10^{-28}$ g.

Planck's constant, $h = 6.547 \times 10^{-27}$ erg $\cdot$ sec.

Avogadro's number, $N_0 = 6.064 \times 10^{23}$ mole^{-1}

Boltzmann constant $k = R/N_0 = 1.3709 \times 10^{-16}$ erg $\cdot$ deg.$^{-1}$

Kinetic energy per molecule at 0°C.

$$E = \tfrac{3}{2}k(273.18) = 5.617 \times 10^{-14} \text{ erg}.$$

Atomic specific heat constant, $\beta = h/k = 4.7757 \times 10^{-11}$ sec $\cdot$ deg.

Mass of hydrogen atom, $m_{\mathrm{H}} = \dfrac{1.008}{N_0} = 1.6618 \times 10^{-24}$ g.

Einstein's constant, $h/e = 1.3725 \times 10^{-17}$ erg $\cdot$ sec $\cdot$ es-unit^{-1}

Energy of V-abs-volt-electron,

$$h\nu = \tfrac{1}{2}m_0v^2 = Ve = 1.591 \times 10^{-12} \text{ erg} \times V.$$

Wave number associated with V-abs-volt-electron,

$$\tilde{\nu}_0 = 8106 \text{ cm.}^{-1} \times V.$$

Wave length associated with V-abs-volt-electron,

$$\lambda_0 = \frac{1}{\nu_0} = 12{,}336 \times 10^{-8} \text{ cm} \cdot V^{-1}.$$

Mechanical equivalent of heat,

$$J_{15} = 4.1852 \text{ abs-joule} \cdot \text{cal}_{15}^{-1}.$$

[1] These values are taken from the publications by R. T. Birge, *Phys. Rev. Suppl.*, **1**, 1 (1929), and *Phys. Rev.*, **40**, 228 (1932).

Energy per mole for 1 abs-volt-electron per molecule,

$$\frac{F}{J_{15}} = 23{,}055 \text{ cal}_{15} \cdot \text{mole}^{-1}.$$

Radius of Bohr orbit,

$$a_0 = \frac{h^2}{4\pi^2 m_0 e^2} = 0.5282 \times 10^{-8} \text{ cm.}$$

Bohr unit of angular momentum,

$$\mathcal{H} = \frac{h}{2\pi} = 1.042 \times 10^{-27} \text{ erg} \cdot \text{sec.}$$

Magnetic moment of 1 Bohr magneton,

$$\mu_B = \frac{he}{4\pi m_0} = 9.175 \times 10^{-21} \text{ erg} \cdot \text{gauss}^{-1}$$

Rydberg constant for infinite nuclear mass,

$$R_\infty = \frac{2\pi^2 e^4 m_0}{h^3 c} = 109{,}737.42 \text{ cm.}^{-1}$$

Rydberg constant for hydrogen, $R_{\text{H}} = 109{,}677.76 \text{ cm.}^{-1}$
Ionization potential for H atom,

$$R_{\text{H}} ch = \frac{2\pi^2 e^4 m}{h^2} = 13.530 \text{ abs-volt.}$$

Schroedinger constant for electron

$$\frac{8\pi^2 m_0}{h^2} = 1.664 \times 10^{27} \text{ g} \cdot \text{ erg}^{-2} \text{ sec.}^{-2}$$

Schroedinger constant for H atom

$$\frac{8\pi^2 m_{\text{H}}}{h^2} = 3.062 \times 10^{30} \text{ g} \cdot \text{ erg}^{-2} \cdot \text{sec.}^{-2}$$

De Broglie wave length associated with V-abs-volt-electron

$$\lambda = \frac{h}{m_0 v} = \frac{h}{\sqrt{2 m_0 V e}} = 12.21 \times 10^{-8} \text{ cm} \cdot V^{-\frac{1}{2}}$$

APPENDIX III

SPECIAL TABLES OF MATHEMATICAL FORMULAS

The following tables contain a number of relations which are of special interest in quantum mechanics calculations. For more complete tables the reader should consult the references given in the last section.

1. Series Expansions. On the basis of the binomial theorem it is shown that for positive integral values of n

$$(x+y)^n = x^n + nx^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 + \cdots$$

$$+ \frac{n(n-1)\cdots(n-k+1)}{k!}x^{n-k}y^k + \cdots$$

$$+ nxy^{n-1} + y^n.$$

For fractional and negative values of n the series converges for $y/x < 1$. If $n > 0$, the series also converges for $|y/x| = 1$.
Special cases:

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \cdots$$

$$(1-x)^{-1} = 1 + x + x^2 + x^3 + \cdots$$

$$(1+x)^{\frac{1}{2}} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \frac{5x^4}{128} + \frac{7x^5}{256} - \frac{21x^6}{1024} + \cdots$$

$$(1+x)^{-\frac{1}{2}} = 1 - \frac{x}{2} + \frac{3x^2}{8} - \frac{5x^3}{16} + \frac{35x^4}{128} - \frac{63x^5}{256} + \frac{231x^6}{1024} - \cdots$$

$$\epsilon = \lim_{n=\infty}\left(1+\frac{1}{n}\right)^n = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots = 2.718282$$

$$\epsilon^{ax} = 1 + \frac{ax}{1!} + \frac{a^2x^2}{2!} + \frac{a^3x^3}{3!} + \cdots$$

$$= b^x, \text{ where } a = \log_\epsilon b = \ln b = 2.30259 \log_{10} b.$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

$$\sinh x = -i \sin (ix) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots$$

$$\cosh x = \cos (ix) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots$$

$$\pi = 3.141593.$$

2. Algebraic Equations. (1) The two roots of the quadratic equation

$$ax^2 + bx + c = 0$$

are given by the relation

$$x = -\frac{b \pm \sqrt{b^2 - 4ac}}{2a}.$$

(2) Given the system of n equations with n unknowns,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= c_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= c_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= c_n \end{aligned}$$

the solution is given by [1]

$$x_k = \frac{1}{D}(c_1D_{1k} + c_2D_{2k} + \dots + c_nD_{nk})$$

$$= \frac{1}{D}\sum_{i=1}^{n} D_{ik},$$

where

$$D = |a_{ik}| = \begin{vmatrix} a_{11}a_{12} \dots a_{1n} \\ a_{21}a_{22} \dots a_{2n} \\ \dots\dots\dots\dots \\ a_{n1}a_{n2} \dots a_{nn} \end{vmatrix},$$

and D_{ik} is a determinant of order $n - 1$ which is obtained thus:

$D_{ik} = (-1)^{i+k} \times$ det. obtained by omitting from D the ith row and kth column.

Hence

$$D = \sum_{k=1}^{n} a_{ik}D_{ik} = \sum_{i=1}^{n} a_{ki}D_{ki}.$$

For $c_1 = c_2 = \dots = c_n = 0$ (*homogeneous* system of equations), $x_1 = x_2 = \dots = x_n = 0$ if $D \neq 0$. For $D = 0$, the system of equations determines the ratios of the unknowns. That is,

$$x_1 : x_2 : \dots : x_n = D_{11} : D_{12} : \dots : D_{1n}.$$

[1] This solution is taken from L. Silberstein's Mathematical Tables.

3. Exponential, Trigonometric, and Hyperbolic Functions.

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y.$$
$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y.$$
$$2i \sin ax = \epsilon^{iax} - \epsilon^{-iax} \text{ (where } i = \sqrt{-1}).$$
$$2 \cos ax = \epsilon^{iax} + \epsilon^{-iax}.$$
$$\epsilon^{iax} = \cos ax + i \sin ax.$$
$$\epsilon^{-iax} = \cos ax - i \sin ax.$$
$$\sinh ax = \frac{1}{2}\left(\epsilon^{ax} - \epsilon^{-ax}\right) = ax + \frac{a^3x^3}{3!} + \frac{a^5x^5}{5!} + \cdots$$
$$\cosh ax = \frac{1}{2}\left(\epsilon^{ax} + \epsilon^{-ax}\right) = 1 + \frac{a^2x^2}{2!} + \frac{a^4x^4}{4!} + \cdots$$
$$\epsilon^{ax} = \cosh ax + \sinh ax.$$
$$\epsilon^{-ax} = \cosh ax - \sinh ax.$$
$$\tanh ax = \frac{\sinh ax}{\cosh ax} = \frac{1 - \epsilon^{-2ax}}{1 + \epsilon^{-2ax}}.$$
$$\cosh^2 ax - \sinh^2 ax = 1.$$

DeMoivre's formula: For integral, fractional, and negative values of n,

$$(\cos\theta + i \sin\theta)^n = \cos n\theta + i \sin n\theta.$$

Given

$$z^2 = x^2 + y^2 = (x + iy)(x - iy),$$

the complex number $(x - iy)$ is the *conjugate* of $(x + iy)$.

$|z| = +\sqrt{x^2 + y^2}$, is known as the *modulus*,
$|z|^2 =$ *norm* of z.

With polar coördinates such that $x = r\cos\theta$, $y = r\sin\theta$,

$$r = |z|; \; z = r(\cos\theta + i\sin\theta) = re^{i\theta}.$$

The complex conjugate of z is designated by $\bar{z}$, and hence

$$\bar{z} = r(\cos\theta - i\sin\theta) = re^{-i\theta}$$
$$\bar{z}z = r^2 = |z|^2.$$

4. Definite Integrals.

$$\int_0^{\frac{\pi}{2}} \sin^2 x dx = \int_0^{\frac{\pi}{2}} \cos^2 x dx = \frac{\pi}{4}.$$
$$\int_0^{\frac{\pi}{2}} \sin^3 x dx = \int_0^{\frac{\pi}{2}} \cos^3 x dx = \frac{2}{3}.$$
$$\int_0^{\frac{\pi}{2}} \sin^4 x dx = \int_0^{\frac{\pi}{2}} \cos^4 x dx = \frac{3}{4} \cdot \frac{\pi}{4}.$$

$$\int_0^{\frac{\pi}{2}} \sin^n x dx = \frac{n-1}{n} \int_0^{\frac{\pi}{2}} \sin^{n-2} x dx \text{ for } n \geqslant 2.$$

$$\int_0^{\frac{\pi}{2}} \cos^n x dx = \frac{n-1}{n} \int_0^{\frac{\pi}{2}} \cos^{n-2} x dx \text{ for } n \geqslant 2.$$

$$\int_0^{\pi} \sin mx \cdot \sin nx dx = \int_0^{\pi} \cos mx \cdot \cos nx dx = 0 \text{ for } m \neq n.$$

$$\int_0^{\frac{\pi}{2}} \sin^m x \cdot \cos^n x dx = \frac{1 \cdot 3 \cdot 5 \ldots (m-1) \cdot 1 \cdot 3 \cdot 5 \ldots (n-1)}{2 \cdot 4 \cdot 6 \ldots (m+n)},$$

where m and n are both *even* integers,

$$= \frac{2 \cdot 4 \cdot 6 \ldots (m-1)}{(n+1)(n+3) \ldots (m+n)},$$

where m is odd.

The *error function* is defined by the integral (that of Gauss)

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x \epsilon^{-x^2} dx.$$

The *exponential integral*

$$Ei(x) = \int_{\infty}^{-x} \frac{\epsilon^{-u}}{u} du.$$

Euler's constant

$$C = \int_0^{\infty} \frac{dx}{x} \left(\frac{1}{1+x^2} - \frac{1}{\epsilon^x} \right) = 0.5772.$$

Evaluation of integrals of the form $J_n = \int_{x_1}^{x_2} x^n \epsilon^{-x} dx.$

$$\text{Since } \frac{d}{dx}(x^n \epsilon^{-x}) = n x^{n-1} \epsilon^{-x} - x^n \epsilon^{-x},$$

$$\int_{x_1}^{x_2} x^n \epsilon^{-x} dx = n \int_{x_1}^{x_2} x^{n-1} \epsilon^{-x} dx - (x_2^n \epsilon^{-x_2} - x_1^n \epsilon^{-x_1}).$$

That is, the value of J_n is expressed in terms of J_{n-1}. Proceeding in this manner it is shown that

$$\int_x^{\infty} x^n \epsilon^{-x} dx = \epsilon^{-x} n! \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^k}{k!} + \ldots + \frac{x^n}{n!} \right).$$

Consequently,

$$\int_0^{\infty} x^n \epsilon^{-x} dx = n!,$$

and

$$\int_0^x x^n \epsilon^{-x} dx = n! - \int_x^{\infty} x^n \epsilon^{-x} dx$$

From these relations the following integrals are derived:

$$\int_0^\infty \epsilon^{-x}dx = 1; \qquad \int_x^\infty \epsilon^{-x}dx = \epsilon^{-x}; \qquad \int_0^x \epsilon^{-x}dx = 1 - \epsilon^{-x}.$$

$$\int_0^\infty x\epsilon^{-x}dx = 1; \qquad \int_x^\infty x\epsilon^{-x}dx = \epsilon^{-x}(1 + x);$$

$$\int_0^x x\epsilon^{-x}dx = 1 - \epsilon^{-x}(1 + x),$$

$$\int_0^\infty x^2\epsilon^{-x}dx = 2; \qquad \int_x^\infty x^2\epsilon^{-x}dx = 2\epsilon^{-x}\left(1 + x + \frac{x^2}{2}\right).$$

$$\int_0^\infty x^3\epsilon^{-x}dx = 6; \qquad \int_x^\infty x^3\epsilon^{-x}dx = 6\epsilon^{-x}\left(1 + x + \frac{x^2}{2} + \frac{x^3}{6}\right).$$

Integrals involving $x^n\epsilon^{-mx}$ may be reduced to the same form thus:

$$\int x^n\epsilon^{-mx}dx = \frac{1}{m^{n+1}}\int (mx)^n\epsilon^{-mx}d(mx) = \frac{1}{m^{n+1}}\int y^n\epsilon^{-y}dy.$$

R. Eisenschitz and F. London, *Z. Physik*, **60**, 491 (1930), give a table of "Exchange" and Coulomb integrals which occur in the treatment of the problem of interaction of two hydrogen atoms.

The *Gamma Function* or *Euler's Integral of the Second Kind* is defined thus:

$$\int_0^\infty x^{n-1}\epsilon^{-x}dx = \Gamma(n).$$

For positive integral values of n

$$\Gamma(n) = \Pi(n-1) = 1 \cdot 2 \ldots (n-1) = (n-1)!$$
$$\Gamma(1) = \Pi(o) = 1; \quad \Gamma(o) = \infty.$$

Integrals which are of special interest in the kinetic theory of gases (see Appendix in J. H. Jeans' "The Dynamical Theory of Gases"):

$$\int_0^\infty x^{2n}\epsilon^{-\lambda x^2}dx = \frac{1 \cdot 3 \cdot 5 \ldots (2n-1)}{2^{n+1}}\sqrt{\frac{\pi}{\lambda^{2n+1}}}$$

$$\int_0^\infty x^{2n+1}\epsilon^{-\lambda x^2}dx = \frac{n!}{2\lambda^{n+1}}.$$

Special cases which are of importance:

$$\int_0^\infty \epsilon^{-\frac{mx^2}{2kT}}dx = \frac{1}{2}\sqrt{\frac{2\pi kT}{m}}$$

$$\int_0^\infty x\epsilon^{-\frac{mx^2}{2kT}}dx = \frac{1}{mkT}$$

$$\int_0^\infty x^2\epsilon^{-\frac{mx^2}{2kT}}dx = \frac{1}{4}\sqrt{\frac{8\pi k^3T^3}{m^3}}$$

$$\int_0^\infty x^3\epsilon^{-\frac{mx^2}{2kT}}dx = \frac{1}{8mk^2T^2}$$

5. References for Numerical Values of Functions and Integrals.

(CH) " Handbook of Chemistry and Physics," Chemical Rubber Publishing Co., Cleveland, Ohio. Annual publication.
(ST) " Smithsonian Physical Tables," eighth revised edition, edited by F. E. Fowle, Smithsonian Institution, Washington (1933).
(MS) " Mathematical Tables," L. Silberstein, G. Bell & Sons, Ltd., London, England.
(MP) " A Short Table of Integrals," B. O. Peirce, Ginn & Co., Boston.
(JE) " Funktionentafeln mit Formeln und Kurven," E. Jahnke und F. Emde, B. G. Teubner, Leipzig and Berlin, 1933.

Function	*Reference for Tables*
$\ln x$, ϵ^x, ϵ^{-x} / $\cosh x$, $\sinh x$	CH; ST; MP; JE
n^{-1}, n^2, n^3, $n^{\frac{1}{2}}$, $n^{\frac{3}{2}}$	CH; ST, (n = 1 to 999)
$n!$, $\log n!$	CH, (n = 1 to 100)
n^4, n^5, n^6, n^7, n^8	CH, (n = 1 to 100)
Erf function	MS, MP
ϵ^{x^2}, ϵ^{-x^2}	ST
$Ei(x)$; $\Gamma(n)$ for n = 1 to 2	ST; JE
Diffusion integral $1 - \frac{2}{\sqrt{\pi}}\int_0^q \epsilon^{-q^2} dq$	ST
Zonal spherical harmonics,	ST, JE
Derived Legendre polynomials,	JE.

APPENDIX IV

SOME FUNDAMENTAL THEOREMS AND DIFFERENTIAL EQUATIONS

Quantum mechanics is a development which has its basis in classical physics. It represents an evolution and not a revolution, as has been imagined by many, especially those who have not taken the trouble to study the logic and technic of the subject. From this point of view the student of quantum mechanics will find it of interest to renew his acquaintanceship with some of the concepts and theorems which should be regarded as the foundations of theoretical physics.

In the following sections an attempt has been made to present a summary, or rather digest, of some of the theorems and differential equations which are of a fundamental nature and of such general scope that they have found application in many diverse fields of physics.

Many of these theorems may be derived either by direct mathematical argument or by application of the concepts of vector analysis. The latter method is the more concise, and, though its symbolism may seem artificial, there is a deep physical significance in the use of vectors.

The presentation given in the following sections follows closely the excellent summary given by R. E. Doherty and E. G. Keller in "Mathematics of Modern Engineering," Vol. I, Chapter III. The bibliography at the end of the chapter gives references to a number of other works on the topic of vector analysis.

1. Vectors and Vector Products. All elementary physical quantities may be classified into two groups: *vectors* and *scalars*. A scalar is that which has magnitude only; a vector has both magnitude and direction. For instance, mass, length, and energy are scalars, but velocity, force, and momentum are examples of vector quantities.

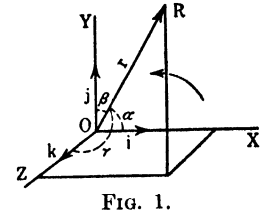


FIG. 1.

A vector is represented, as shown in Fig. 1, by a line drawn in the direction in which the magnitude is effective, and an arrowhead is used to indicate the sense. To distinguish between them, vectors are usually printed in **bold-face** type, and scalars in *italics*.

The vector **R** has the magnitude r, and the direction indicated by the arrow. The convention regarding sign is as follows: The positive direction of the z-axis is taken to be that in which a right-handed screw advances as it is rotated in the xy-plane from the positive direction of the x-axis to that of the y-axis, as indicated in Fig. 1 by the circular arrow in the xy-plane.

The direction of a vector is usually indicated by the three direction cosines (see Fig. 1)

$$\left.\begin{aligned} l &= \cos(x, r) = \cos\alpha \\ m &= \cos(y, r) = \cos\beta \\ n &= \cos(z, r) = \cos\gamma \end{aligned}\right\}. \tag{1}$$

Hence,

$$l^2 + m^2 + n^2 = 1. \tag{2}$$

If l_1, m_1, n_1 and l_2, m_2, n_2 are the direction cosines of two vectors **R** and **Q**, drawn from the same origin, the angle θ between the two lines is given by

$$\cos\theta = l_1l_2 + m_1m_2 + n_1n_2. \tag{3}$$

It follows that if the two vectors are at right angles to each other,

$$l_1l_2 + m_1m_2 + n_1n_2 = 0. \tag{4}$$

A unit vector is one whose magnitude is unity, and the three unit vectors **i**, **j**, **k** correspond to vectors of unit length directed along the three rectangular coördinate axes x, y, z, respectively, as shown in Fig. 1. Any vector $\mathbf{R}_1$ is therefore described by the three component vectors thus:

$$\mathbf{R}_1 = \mathbf{i}x_1 + \mathbf{j}y_1 + \mathbf{k}z_1, \tag{5}$$

where r_1 the magnitude of the vector **R** is given by the relation

$$r_1^2 = x_1^2 + y_1^2 + z_1^2.$$

If we have another vector

$$\mathbf{R}_2 = \mathbf{i}x_2 + \mathbf{j}y_2 + \mathbf{k}z_2,$$

making an angle θ with the vector $\mathbf{R}_1$, the resultant vector **Q** is given by the law of parallelogram of velocities

$$|Q|^2 = |R_1|^2 + |R_2|^2 + 2|R_1||R_2|\cos\theta, \tag{6}$$

where the vertical lines indicate that the absolute value of the vector is to be used.

The *scalar* or *dot product* of two vectors is a *scalar* and is written in the form $\mathbf{R}_1 \cdot \mathbf{R}_2$. It is defined by the relation

$$\mathbf{R}_1 \cdot \mathbf{R}_2 = (\mathbf{i}x_1 + \mathbf{j}y_1 + \mathbf{k}z_1) \cdot (\mathbf{i}x_2 + \mathbf{j}y_2 + \mathbf{k}z_2). \tag{7}$$

Since
$$\left.\begin{aligned} \mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} &= 1 \\ \text{and} \quad \mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} &= 0 \end{aligned}\right\}, \tag{8}$$

it follows that
$$\mathbf{R}_1 \cdot \mathbf{R}_2 = \mathbf{R}_2 \cdot \mathbf{R}_1 = x_1x_2 + y_1y_2 + z_1z_2. \tag{9}$$

That is, the order of " multiplication " is unimportant.

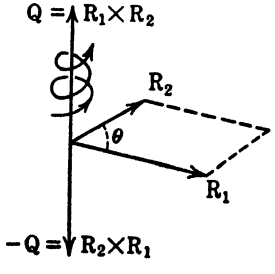


FIG. 2.

The *vector* or *cross product* of two vectors is a *vector* perpendicular to the plane containing the two vectors and is regarded as positive if the product is taken in the same direction as that of the rotation of a right-handed screw. Thus, in Fig. 2

$$\mathbf{Q} = \mathbf{R}_1 \times \mathbf{R}_2 = -\mathbf{R}_2 \times \mathbf{R}_1.$$

That is, the order of " multiplication " is significant, and the product is said to be *non-commutative*. The absolute magnitude of the vector **Q** is equal to the area of the parallelogram of which $\mathbf{R}_1$ and $\mathbf{R}_2$ are the sides. That is,

$$|Q| = |R_1| \times |R_2| \sin\theta. \tag{10}$$

It follows that
$$\left.\begin{aligned} \mathbf{i}\times\mathbf{i} = \mathbf{j}\times\mathbf{j} = \mathbf{k}\times\mathbf{k} &= 0 \\ \mathbf{j}\times\mathbf{k} = \mathbf{i} &= -\mathbf{k}\times\mathbf{j} \\ \mathbf{k}\times\mathbf{i} = \mathbf{j} &= -\mathbf{i}\times\mathbf{k} \\ \mathbf{i}\times\mathbf{j} = \mathbf{k} &= -\mathbf{j}\times\mathbf{i} \end{aligned}\right\}. \tag{11}$$

It is important to note the cyclical order of the factors in the last three equations.

2. Line and Surface Integrals of Vectors. (*a*) *Line Intergrals.* The work done by a force F on a particle in moving a distance dr is given by $F \cos \theta dr$, where θ is the angle between the direction of F and that of dr. (In Fig. 3 dr is directed along the tangent to the curve.) In vector notation, work done in traversing the distance AB along the curve is given by

$$\int_A^B \mathbf{F} \cdot d\mathbf{r} = \int_A^B F \cos \theta dr$$

$$= \int_A^B (\mathbf{i}F_x + \mathbf{j}F_y + \mathbf{k}F_z) \cdot (\mathbf{i}dx + \mathbf{j}dy + \mathbf{k}dz)$$

$$= \int_A^B (F_x dx + F_y dy + F_z dz). \tag{12}$$

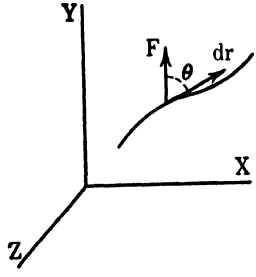


FIG. 3.

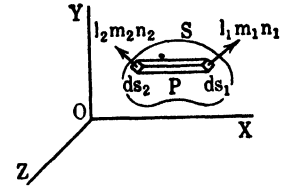


FIG. 4.

(*b*) *Surface Integrals.* Consider a region in which there is a potential energy function. Let S designate a closed surface drawn in this region (see Fig. 4), and consider a prism P, in this region, of cross section $dydz$, the sides of which are parallel to the OX-axis. Let dS_1 and dS_2 denote the intercepts of this prism on the surface S, and let $l_1m_1n_1$ and $l_2m_2n_2$ denote the direction cosines of the normal to these two elements of surface respectively.

Then $dydz = l_1 dS_1 = -l_2 dS_2$, and $F_x dydz = F_1 l_1 dS_1 - F_2 l_2 dS_2$ denotes the *net flux* through the prism P in the outward direction, where F_1 and F_2 designate the values of F_x at x_1, y, z and x_2, y, z respectively.

More generally, if $\mathbf{n}$ designates the unit distance along the normal to any element dS of the surface, this element can be represented by a *vector*

$$d\mathbf{S} = \mathbf{n}dS, \tag{13}$$

and the integral

$$\int_S \mathbf{F} \cdot \mathbf{n}dS = \int_S \mathbf{F} \cdot d\mathbf{S} \tag{14}$$

is the surface integral of $\mathbf{F}$ over the surface S.

The surface integral thus defined evidently corresponds to the total *flux* of $\mathbf{F}$ through the surface S. If $\mathbf{F}$ is the intensity of heat flow (e.g., in calories) per unit time, per unit area *perpendicular* to the direction of flow, the surface integral represents the total heat flow through the surface per unit time.

Similarly, $\mathbf{F}$ may represent a velocity vector, e.g., the mass of fluid crossing unit area per unit time. Then the surface integral represents the total flow through the surface, per unit time.

3. Derivative of a Vector. Let us consider the vector $\mathbf{r}$ of magnitude r, where

$$\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z, \tag{15}$$

and x, y, and z are each functions of t.

Then $\dot{r}$ defines the magnitude of a vector $\dot{\mathbf{r}}$, and we can express this vector in terms of the components thus:

$$\dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt} = \mathbf{i} \cdot \frac{dx}{dt} + \mathbf{j} \cdot \frac{dy}{dt} + \mathbf{k} \cdot \frac{dz}{dt}. \tag{16}$$

Similarly, we can express the acceleration in the direction of r as a vector by the relation

$$\frac{d^2\mathbf{r}}{dt^2} = \mathbf{i} \cdot \frac{d^2x}{dt^2} + \mathbf{j} \cdot \frac{d^2y}{dt^2} + \mathbf{k} \cdot \frac{d^2z}{dt^2}. \tag{17}$$

It is readily shown that

$$\frac{d}{dt}(\mathbf{R}_1 \cdot \mathbf{R}_2) = \frac{d\mathbf{R}_1}{dt} \cdot \mathbf{R}_2 + \mathbf{R}_1 \cdot \frac{d\mathbf{R}_2}{dt}, \tag{18}$$

$$\frac{d}{dt}(\mathbf{R}_1 \times \mathbf{R}_2) = \frac{d\mathbf{R}_1}{dt} \times \mathbf{R}_2 + \mathbf{R}_1 \times \frac{d\mathbf{R}_2}{dt}, \tag{19}$$

and that if $\mathbf{R} = \mathbf{R}(x, y, z)$, then

$$\int_\tau \mathbf{R} d\tau = \mathbf{i}\int R_x d\tau + \mathbf{j}\int R_y d\tau + \mathbf{k}\int R_z d\tau, \tag{20}$$

where $d\tau = dxdydz$.

4. The Gradient. Let $V(x, y, z)$ be a scalar point function. For instance, V may designate the potential energy function, or the temperature at any point in a given region. It is assumed that V is a continuous and singly valued function of the coördinates.

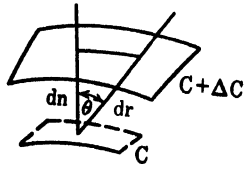


FIG. 5.

For any value $V = C$, a constant, this function will be represented by a surface, designated an *equipotential*. Let us consider two adjacent surfaces defined by the equations $V = C$ and $V + \Delta V = C + \Delta C$ (see Fig. 5). The derivative of V along the normal is dV/dn, and this is greater than the rate of change of V in any other direction, dV/dr, since dn is less than dr.

Evidently dV/dn is a vector, and it is known as the *gradient* of V. This is usually designated thus:

$$\nabla \cdot V = \text{grad } V = \mathbf{n} \cdot \frac{dV}{dn}, \tag{21}$$

where $\mathbf{n}$ is a unit vector in the direction of the normal.

Since

$$\left(\frac{dV}{dn}\right)^2 = \left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2,$$

it follows that we can also write

$$\nabla \cdot V = \mathbf{i} \cdot \frac{\partial V}{\partial x} + \mathbf{j} \cdot \frac{\partial V}{\partial y} + \mathbf{k} \cdot \frac{\partial V}{\partial z}, \tag{22}$$

where the symbol ∇, called " del," is a *vector operator* defined by the relation

$$\nabla = \mathbf{i} \cdot \frac{\partial}{\partial x} + \mathbf{j} \cdot \frac{\partial}{\partial y} + \mathbf{k} \cdot \frac{\partial}{\partial z}. \tag{23}$$

From this we deduce the operator

$$\nabla \cdot \nabla = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \tag{24}$$

which is designated " del squared " and is also known as the *Laplacian.*

As an application of the concept of gradient let us consider the flow of heat through a medium. The flow occurs in the direction of greatest rate of decrease of temperature, and is given by the vector relation

$$\mathbf{q} = -k \cdot \nabla T, \tag{25}$$

where q is the heat flux, ∇T is the gradient of temperature, and k is defined as the coefficient of heat conductivity.

It will be observed that the dot product of the vector operator ∇ and a scalar function is a *vector*, which is designated the gradient of the scalar function.

5. Divergence. The result of operating with ∇ on a *vector* function $\mathbf{F}(x, y, z)$ is a *scalar* function which is designated the *divergence* of F. From equation (23) we have the relation

$$\nabla \cdot \mathbf{F} = \mathbf{i} \cdot \frac{\partial \mathbf{F}}{\partial x} + \mathbf{j} \cdot \frac{\partial \mathbf{F}}{\partial y} + \mathbf{k} \cdot \frac{\partial \mathbf{F}}{\partial z}. \tag{26}$$

Let F_x, F_y, and F_z denote the rectangular component of F. Then

$$\mathbf{F} = \mathbf{i}F_x + \mathbf{j}F_y + \mathbf{k}F_z.$$

Hence,

$$\begin{aligned} \nabla \cdot \mathbf{F} &= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \cdot (\mathbf{i}F_x + \mathbf{j}F_y + \mathbf{k}F_z) \\ &= \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \\ &= \operatorname{div} \mathbf{F}. \end{aligned} \tag{27}$$

If we regard F as a flux, that is, number of lines per unit area, the term divergence means the number of lines which spread out per unit volume. This interpretation is illustrated best by the application of the concept in the derivation of the equation of continuity.

6. The Equation of Continuity. Let us consider a rectangular element of volume $dxdydz = d\tau$, at x, y, z, in a fluid of density $\rho = \rho(x, y, z)$. Let v denote the velocity of the fluid at the given point. The mass of fluid M, flowing through a unit area, per unit time, is given by the mass of a prism of fluid of unit area and length v. That is

$$\mathbf{M} = \rho \mathbf{v}, \tag{28}$$

where $\mathbf{M}$ and $\mathbf{v}$ are vectors. Let $\mathbf{M}_x$, $\mathbf{M}_y$, and $\mathbf{M}_z$ denote the rectangular components of the vector $\mathbf{M}$.

The rate at which fluid enters the area $dydz$ in the positive direction of the x-axis, at x, y, z, is

$$\rho v dydz = M_x dydz.$$

The rate at which fluid leaves the opposite area at $x + dx, y, z$ is

$$\left\{\rho v_x + \frac{\partial}{\partial x}(\rho v_x)dx\right\} dydz = \left(M_x + \frac{\partial M_x}{\partial x}dx\right)dydz.$$

Hence the net increase in the mass of fluid inside the element $d\tau$, due to the vector component $\mathbf{M}_x$, is

$$-\frac{\partial M_x}{\partial x}d\tau.$$

Adding the net increments per unit time due to the inflow and outflow through the other four sides of the element $d\tau$, we obtain the relation

$$\frac{\partial \rho}{\partial t} = -\left(\frac{\partial M_x}{\partial x} + \frac{\partial M_y}{\partial y} + \frac{\partial M_z}{\partial z}\right). \tag{29}$$

That is,

$$\nabla \cdot (\rho \mathbf{v}) = \nabla \cdot \mathbf{M} = \operatorname{div} \mathbf{M} = -\frac{\partial \rho}{\partial t}. \tag{30a}$$

This is known as the *equation of continuity*. If ρ is constant throughout the fluid, this equation assumes the form

$$\frac{\partial \rho}{\partial t} + \rho \operatorname{div} \mathbf{v} = 0. \tag{30b}$$

If the fluid is incompressible, $\partial\rho/\partial t = 0$, and hence

$$\nabla \cdot \mathbf{M} = 0.$$

"The name divergence," as L. Page remarks, "originated in this interpretation of $\nabla \cdot \mathbf{M}$. For since $-\nabla \cdot \mathbf{M}$ represents the excess of the inward over the outward flow, or the *convergence* of the fluid, so $\nabla \cdot \mathbf{M}$ represents the excess of the outward over the inward flow, or the *divergence* of the fluid."[1]

7. Gauss' Theorem. Let $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ be the components of a vector $\mathbf{F}$, a continuous function of the coördinates, and let S designate a closed surface enclosing a region of volume V.

Gauss' theorem states that

$$\int_V \left(\frac{\partial \mathbf{X}}{\partial x} + \frac{\partial \mathbf{Y}}{\partial y} + \frac{\partial \mathbf{Z}}{\partial z}\right)d\tau = \int_S (l\mathbf{X} + m\mathbf{Y} + n\mathbf{Z})dS, \tag{31}$$

where l, m, n are the direction cosines of the normal drawn *outwards* to the surface element dS.

Since

$$\int_{x_2}^{x_1} \frac{\partial \mathbf{X}}{\partial x}dx = \mathbf{X}_1 - \mathbf{X}_2,$$

[1] "Introduction to Theoretical Physics," p. 22.

it follows that

$$\int dz \int dy \int \frac{\partial \mathbf{X}}{\partial x} dx = \int dz \int dy (\mathbf{X}_1 - \mathbf{X}_2),$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ are the values of $\mathbf{X}$ at the two elements dS_1 and dS_2 in which the prism of cross section $dydz$ intersects the surface S (see Fig. 4).

But

$$dydz = l_1 dS_1 = -l_2 dS_2.$$

Hence

$$(\mathbf{X}_1 - \mathbf{X}_2) dydz = l_1 \mathbf{X}_1 dS_1 - l_2 \mathbf{X}_2 dS_2.$$

Integrating through the elementary parallel prisms and adding, we obtain the relation

$$\int_V \frac{\partial \mathbf{X}}{\partial x} d\tau = \int_S l \mathbf{X} dS.$$

In a similar manner it is shown that

$$\int_V \frac{\partial \mathbf{Y}}{\partial y} d\tau = \int_S m \mathbf{Y} dS, \quad \text{and} \quad \int_V \frac{\partial \mathbf{Z}}{\partial z} d\tau = \int_S n \mathbf{Z} dS,$$

from which equation (31) follows.

From equation (27) it follows that (31) may be written in the form

$$\int_V \nabla \cdot \mathbf{F} d\tau = \int_S (l \mathbf{F}_x + m \mathbf{F}_y + n \mathbf{F}_z) dS.$$

But the integrand on the right-hand side is equal to $F \cos \theta$, where θ is the angle between the direction of F and the normal at the element dS. Hence,

$$\int_V \nabla \cdot \mathbf{F} d\tau = \int_S \mathbf{F} \cdot \mathbf{n} dS = \int_S \mathbf{F} \cdot d\mathbf{S}. \tag{32a}$$

If $\mathbf{F}$ is the derivative of a scalar point function V, which together with its derivatives are finite and continuous over the configuration space, then

$$F_x = -\frac{\partial V}{\partial x}, \; F_y = -\frac{\partial V}{\partial y}, \; F_z = -\frac{\partial V}{\partial z},$$

and

$$F \cos \theta = -\frac{\partial V}{\partial n}.$$

Hence, equation (32a) may be written in the form

$$\int_V \nabla \cdot \mathbf{F} d\tau = \int_V (\operatorname{div} \mathbf{F}) d\tau = -\int \frac{\partial V}{\partial n} dS, \tag{32b}$$

or as

$$\int_V \nabla^2 V d\tau = \int_S \frac{\partial V}{\partial n} dS = \int_S (\operatorname{grad} V) dS. \tag{32c}$$

Gauss' theorem (to be distinguished from Gauss' law, which is discussed in a subsequent section) states a relation between the integral throughout a region V

of the configuration space and the integral taken over a surface S enclosing the given region. It is one of the most fundamental theorems in theoretical physics.

8. Cross Product of Vectors. An important application of such a product is the theorem designated as *Stokes' Theorem*,[2] which states a relation between a surface and line integral (see section 2) in the form

$$\int_S \nabla \times \mathbf{F} \cdot d\mathbf{S} = \int_C \mathbf{F} \cdot d\mathbf{r}, \tag{33}$$

where the second integral is taken along the entire curve C, enclosing the surface S. The operator $\nabla \times \mathbf{F}$ is known as the *curl of the vector function* $\mathbf{F}$.

From the rules for forming the cross product of two vectors it follows that

$$\nabla \times \mathbf{F} = \mathbf{i}\left(\frac{\partial Fz}{\partial y} - \frac{\partial F_y}{\partial z}\right) + \mathbf{j}\left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right) + \mathbf{k}\left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right) \tag{34}$$

$$= \text{curl } \mathbf{F}.$$

The curl of $\mathbf{F}$ thus has the components defined by the bracket terms in the last equation. This operator is of special significance in electromagnetic theory, but is also used extensively in other fields of physics.

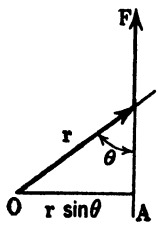


FIG. 6.

Another illustration of the cross product of vectors is obtained from mechanics. The moment of a force $\mathbf{F}$ about a point (see Fig. 6) is defined by the magnitude $F \times r \sin\theta$. This corresponds to a vector

$$\mathbf{M} = \mathbf{r} \times \mathbf{F} \tag{35}$$

of which the direction, as follows from Fig. 2, would be perpendicular to the paper and pointing directly up.

Resolving the vectors $\mathbf{r}$ and $\mathbf{F}$ into their rectangular components,

$$\mathbf{M} = \mathbf{r} \times \mathbf{F} = (\mathbf{i}x + \mathbf{j}y + \mathbf{k}z) \times (\mathbf{i}F_x + \mathbf{j}F_y + \mathbf{k}F_z).$$

Applying the relations for cross products of unit vectors given in (11), we obtain the relation

$$\mathbf{M} = \mathbf{i}M_x + \mathbf{j}M_y + \mathbf{k}M_z,$$

where

$$\left.\begin{aligned} M_x &= (yF_z - zF_y) \\ M_y &= (zF_x - xF_z) \\ M_z &= (xF_y - yF_x) \end{aligned}\right\}. \tag{36}$$

9. Equation for Flow of Heat. In section 3 it was pointed out that the heat flux through a body is given by the relation

$$\mathbf{q} = -k\nabla T.$$

If there are no sources of heat within the surface, the heat flux through the surface must be equal to the rate at which heat leaves the body. Hence

$$\int_V -\frac{\partial T}{\partial t} c\rho d\tau = \int_S -k\nabla T dS, \tag{37}$$

where c = heat capacity per unit mass,
ρ = density.

[2] For proof of this theorem see L. Page, *op. cit.*, p. 31.

The right-hand side of (37) can be written in the form

$$-\int_S k\nabla T dS = \int_S \mathbf{n}\cdot \mathbf{q} dS, \tag{38}$$

where $\mathbf{n}$ is the unit vector normal to the surface at the element dS.

But in consequence of Gauss' theorem, equation (32a),

$$\int_S \mathbf{n}\cdot \mathbf{q} dS = \int_V \nabla\cdot \mathbf{q} d\tau = -\int \nabla\cdot k\nabla T d\tau.$$

Hence we derive from equation (37) the partial differential equation

$$-\frac{\partial T}{\partial t} = \frac{1}{c\rho}\nabla\cdot k\nabla T. \tag{39}$$

If k is assumed constant, this equation becomes

$$-\frac{\partial T}{\partial t} = \frac{k}{c\rho}\nabla^2 T. \tag{40}$$

While equation (39) is the general differential equation for heat flow, equation (40) is the form usually adopted. The solutions of this equation give T as a function of both t and the space coordinates.

Equation (40) can be applied to other types of " flow " as well as heat flow. Instead of a gradient of T, we can have, in a solution, a gradient of C, the concentration of solute. The resulting form of equation (40) is

$$-\frac{\partial C}{\partial t} = D\cdot \nabla^2 C, \tag{41}$$

where D is the coefficient of diffusion. Equation (41) is known as Fick's equation.

10. Equation for Wave Motion. The differential equation for the propagation of a wave motion may be deduced directly, as in section 2.4, or by application of vector analysis. In terms of rectangular coördinates the equation has the form

$$\frac{\partial^2\phi}{\partial t^2} = v^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\phi, \tag{42}$$

where ϕ is the amplitude and v is a constant which is known as the phase velocity.

In order to integrate this equation we use the method of separation of the variables and postulate, as in section 2.4, that we can represent $\phi(t, x, y, z)$ as a product of four functions in the form

$$\phi(t, x, y, z) = g(t)\cdot X(x)\cdot Y(y)\cdot Z(z).$$

If this is a solution, then we must obtain the relation

$$\frac{1}{v^2}\cdot\frac{\ddot{g}(t)}{g(t)} = \frac{1}{X}\cdot\frac{\partial^2 X}{\partial x^2}+\frac{1}{Y}\cdot\frac{\partial^2 Y}{\partial y^2}+\frac{1}{Z}\cdot\frac{\partial^2 Z}{\partial z^2} = -m^2,$$

where m^2 is a constant.

This leads to the solutions

$$\begin{aligned} g(t) &= A_1\epsilon^{imvt}+A_2\epsilon^{-imvt},\\ X(x) &= B_1\epsilon^{i\alpha x}+B_2\epsilon^{-i\alpha x},\\ Y(y) &= C_1\epsilon^{i\beta y}+C_2\epsilon^{-i\beta y},\\ Z(z) &= D_1\epsilon^{i\gamma z}+D_2\epsilon^{-i\gamma z}, \end{aligned}$$

where $\alpha^2+\beta^2+\gamma^2 = m^2$.

Evidently, mv has the dimensions of a frequency ν, and therefore m must have the dimensions of a reciprocal length L which is fixed by the boundary conditions of the problem. That is, $m = n\pi/L$, where n will be an even or odd integer, and $2L/n = \lambda$, a wave length which has three components, one for each of the axes of coördinates.

11. Gauss' Law. Let us consider a surface S drawn about the point P (see Fig. 7) at which is located a positive charge e. With P as vertex draw a cone of solid

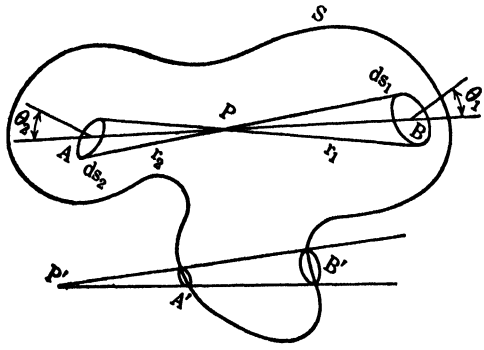


FIG. 7.

angle $d\omega$ and let dS_1 and dS_2 designate the intercepting surface elements. Let θ_1 and θ_2 designate the angles which the axis of the cones makes with the normals to the surface element.

Then,

$$dS_1 \cos\theta_1 = r_1^2 d\omega$$
$$dS_2 \cos\theta_2 = r_2^2 d\omega,$$

so that

$$d\omega = \frac{dS_1 \cos\theta_1}{r_1^2} = \frac{dS_2 \cos\theta_2}{r_2^2}. \tag{43}$$

But the force $\mathbf{F}$ on a unit charge, directed outwards along the normal, is given by $e/r_1^2 \cos\theta_1$ at dS_1 and by $e/r_2^2 \cos\theta_2$ at dS_2. Hence

$$F_2 dS_2 + F_1 dS_1 = \frac{e}{r_1} \cos\theta_1 dS_1 + \frac{e}{r_2^2} \cos\theta_2 dS_2$$
$$= 2e \cdot d\omega.$$

If now we continue to divide up the surface into elements by an infinite number of cones, and add up the contributions to the total flux through the surface, we shall obtain the relation

$$\int_S F \cos\theta dS = \int_S \mathbf{F} \cdot \mathbf{n} dS = \int_S \mathbf{F} \cdot d\mathbf{S} = 2e \int_0^{2\pi} d\omega = 4\pi e. \tag{44a}$$

This result is evidently valid if instead of one charge located at P we have a number of charges distributed inside the surface S. In that case, if $\rho = \rho(x, y, z)$ designates the distribution function for the positive charge in each element of volume, equation (44a) becomes

$$\int_S \mathbf{F} \cdot d\mathbf{S} = 4\pi \int_V \rho d\tau. \tag{44b}$$

It is evident that, if the surface S is taken *outside* the point P' at which the charge is located (see Fig. 7), then the flux through the element dS_2 (located at A') is *equal and opposite* to that through the element dS_1 (located at B'). Hence equations (44*a*) and (44*b*) assume the form

$$\int_S \mathbf{F} \cdot d\mathbf{S} = 0. \tag{45}$$

Equations (44) and (45) are statements of *Gauss' Law.*

12. Equations of Poisson and Laplace. Let us now combine equations (44*b*) and (45) with Gauss' theorem which was derived in section 7 in the form

$$\int_V \nabla \cdot \mathbf{F} d\tau = \int_S \mathbf{F} \cdot d\mathbf{S}. \tag{32}$$

We thus obtain the relation

$$\int \nabla \cdot \mathbf{F} d\tau = 4\pi \int_V \rho d\tau,$$

which, in consequence of (27), becomes

$$\int_V \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) d\tau = 4\pi \int_V \rho d\tau.$$

That is,

$$\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = 4\pi\rho. \tag{46}$$

Now, if the force components are derived from a potential energy function $V = V(x, y, z)$, then

$$F_x = -\frac{\partial V}{\partial x}; \; F_y = -\frac{\partial V}{\partial y}; \; F_z = -\frac{\partial V}{\partial z};$$

and therefore equation (46) assumes the form

$$\nabla^2 V = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) V = -4\pi\rho. \tag{47}$$

Equation (47) is known as *Poisson's equation.*

If $\rho = 0$, that is, if the potential energy function is taken over a region in which there is no electric charge, equation (47) assumes the form

$$\nabla^2 V = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) V = 0, \tag{48}$$

which is known as *Laplace's equation.*

13. Transformation of Coördinates. In many cases it is more convenient to use a coördinate system different from that involving Cartesian coördinates. By application of Gauss' theorem it is possible to derive the form of the operator for such a system.

Let

$$\left.\begin{aligned} q_1 &= q_1(x, y, z) \\ q_2 &= q_2(x, y, z) \\ q_3 &= q_3(x, y, z) \end{aligned}\right\}$$

denote the generalized curvilinear coördinates as functions of the Cartesian coördinates. In the following considerations it will be assumed that the q-coördinates are of such a nature that three surfaces q_1 = constant, q_2 = constant, and q_3 = constant, intersect *orthogonally*. Such a system is of the type designated " orthogonal curvilinear."

" In Cartesian coördinates the three families of surfaces are families of planes parallel to the different coördinate planes. In polars we have a family of concentric spheres, a family of cones with the same axis and vertex, and a family of planes intersecting in the one straight line. In cylindricals we have a family of coaxial cylinders, a family of planes intersecting in the axis of the cylinders, and a family of planes at right angles to the axis of the cylinders. In elliptic coördinates we have three families of confocal conicoids, ellipsoids, hyperboloids of one sheet and hyperboloids of two sheets."[3]

Let us consider the element of volume bounded by the planes

$$q_1 \pm \frac{dq_1}{2},\ q_2 \pm \frac{dq_2}{2};\ q_3 \pm \frac{dq_3}{2}.$$

This will be approximately rectangular, and the length of the side between the surfaces $q_1 - dq_1/2$ and $q_1 + dq_1/2$ is $a_1 dq_1$, where a_1 is a function of q_1, q_2, and q_3. Similarly, let $a_2 dq_2$ and $a_3 dq_3$ denote the lengths of the other sides where a_2 and a_3 are each functions of the three generalized coördinates.

Let Q_1, Q_2, and Q_3 designate the components of force normal to the surfaces q_1, q_2, and q_3 respectively.

Then the area of the surface at q_1 is $a_2 a_3 dq_2 dq_3$, and the flux of force through this area is $Q_1 a_2 a_3 dq_2 dq_3$. The difference between the flux at $q_1 + dq_1/2$ and that at $q_1 - dq_1/2$ is given by

$$\frac{\partial}{\partial q_1}(Q_1 a_2 a_3) dq_1 dq_2 dq_3.$$

Adding up the net amounts of flux through the other two pairs of surfaces, and applying Gauss' law, the result obtained is

$$\left\{\frac{\partial}{\partial q_1}(Q_1 a_2 a_3) + \frac{\partial}{\partial q_2}(Q_2 a_1 a_3) + \frac{\partial}{\partial q_3}(Q_3 a_1 a_2)\right\} dq_1 dq_2 dq_3 = 4\pi\rho a_1 a_2 a_3 dq_1 dq_2 dq_3.$$

That is,

$$\frac{1}{a_1 a_2 a_3}\left\{\frac{\partial}{\partial q_1}(Q_1 a_2 a_3) + \frac{\partial}{\partial q_2}(Q_2 a_1 a_3) + \frac{\partial}{\partial q_3}(Q_3 a_1 a_2)\right\} - 4\pi\rho = 0, \qquad (49)$$

which is a more generalized form of equation (46).

Let $V = V(q_1 q_2 q_3)$ denote the potential function. Then

$$Q_1 = -\frac{\partial V}{a_1 \partial q_1},\quad Q_2 = -\frac{\partial V}{a_2 \partial q_2},\quad Q_3 = -\frac{\partial V}{a_3 \partial q_3};$$

and equation (49) becomes

$$\frac{\partial}{\partial q_1}\left(\frac{a_2 a_3}{a_1}\frac{\partial V}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{a_1 a_3}{a_2}\frac{\partial V}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{a_1 a_2}{a_3}\frac{\partial V}{\partial q_3}\right) + 4\pi\rho a_1 a_2 a_3 dq_1 dq_2 dq_3 = 0. \qquad (50)$$

[3] R. A. Houstoun, " An Introduction to Mathematical Physics," p. 21, on whose remarks this section is based.

Comparing the last equation with equation (47), it follows that

$$\nabla^2 V = \frac{1}{a_1 a_2 a_3} \sum \frac{\partial}{\partial q_1}\left(\frac{a_2 a_3}{a_1}\frac{\partial V}{\partial q_1}\right). \tag{51}$$

The element of volume is evidently

$$dxdydz = a_1 a_2 a_3 dq_1 dq_2 dq_3, \tag{52}$$

and the element of distance dS is given by the relation

$$(ds)^2 = a_1^2(dq_1)^2 + a_2^2(dq_2)^2 + a_3^2(dq_3)^2. \tag{53}$$

The product $a_1a_2a_3$ is known as the *discriminant* of the transformation.[4]

Illustrations of transformations to cylindrical spherical polar and confocal elliptic coördinates have been given in different sections in the preceding chapters.

14. Green's Theorem. Equations (31) and (44*b*) lead to the relation

$$\int_S (l\mathrm{X} + m\mathrm{Y} + n\mathrm{Z})dS = -\int_V \left(\frac{\partial \mathrm{X}}{\partial x} + \frac{\partial \mathrm{Y}}{\partial y} + \frac{\partial \mathrm{Z}}{\partial z}\right) d\tau = -4\pi \int \rho d\tau. \tag{54}$$

In this equation X, Y, and Z may be written as the force components derived from a potential energy function V, such that

$$X = -\frac{\partial V}{\partial x},\ Y = -\frac{\partial V}{\partial y},\ Z = -\frac{\partial V}{\partial z}.$$

Let

$$X = \frac{\Phi \partial \Psi}{\partial x},$$

$$Y = \frac{\Phi \partial \Psi}{\partial y},$$

$$Z = \frac{\Phi \partial \Psi}{\partial z},$$

where Φ and Ψ are scalar point-functions which with their derivatives are uniform, and continuous in the configuration space. From the equations in section 7 it follows that

$$l\mathrm{X} + m\mathrm{Y} + n\mathrm{Z} = \Phi\left(l\frac{\partial \Psi}{\partial x} + m\frac{\partial \Psi}{\partial y} + n\frac{\partial \Psi}{\partial z}\right) = \Phi\frac{\partial \Psi}{\partial n}.$$

Also

$$\frac{\partial \mathrm{X}}{\partial x} + \frac{\partial \mathrm{Y}}{\partial y} + \frac{\partial \mathrm{Z}}{\partial z} = \frac{\partial}{\partial x}\left(\Phi\frac{\partial \Psi}{\partial x}\right) + \frac{\partial}{\partial y}\left(\Phi\frac{\partial \Psi}{\partial y}\right) + \frac{\partial}{\partial z}\left(\Phi\frac{\partial \Psi}{\partial z}\right)$$

$$= \sum \frac{\partial \Phi}{\partial x}\frac{\partial \Psi}{\partial x} + \Phi \sum \frac{\partial^2 \Psi}{\partial x^2}.$$

Hence

$$\int_V \left(\Phi \nabla^2 \Psi + \sum \frac{\partial \Phi}{\partial x}\cdot\frac{\partial \Psi}{\partial x}\right) d\tau = -\int_S \Phi \frac{\partial \Psi}{\partial n}\cdot dS. \tag{55}$$

[4] It should be observed that the magnitudes designated by a_1, a_2, and a_3 correspond to the *squares* of the magnitudes designated by the same letters in equation (4.22) and in section 6.2.

This is one form of Green's theorem. A second form, which is more generally used, is derived from (55) as follows: If, in equation (55), Φ and Ψ are interchanged, the result is

$$\int_V \left(\Psi\nabla^2\Phi + \sum \frac{\partial\Psi}{\partial x}\cdot\frac{\partial\Phi}{\partial x}\right) d\tau = -\int_S \Psi \frac{\partial\Phi}{\partial n} dS. \tag{56}$$

Now let us subtract (56) from (55). The result is the second form of Green's theorem

$$\int_V (\Phi\nabla^2\Psi - \Psi\nabla^2\Phi) d\tau = -\int_S \left(\Phi\frac{\partial\Psi}{\partial n} - \Psi\frac{\partial\Phi}{\partial n}\right) dS. \tag{57}$$

Because of equation (21) this may be written in the form

$$\int_V (\Phi\nabla^2\Psi - \Psi\nabla^2\Phi) d\tau = -\int_S (\Phi \operatorname{grad}\Psi - \Psi \operatorname{grad}\Phi) dS, \tag{58}$$

which is extremely useful in quantum mechanics.

15. Proof that Solutions of S. Equation Form an Orthogonal System of Functions. In the S. equation

$$-\frac{1}{\alpha^2}\nabla^2\phi + V\phi = E\phi,$$

let ϕ_n and ϕ_m designate two eigenfunctions corresponding to the eigenvalues E_n and E_m respectively. Then

$$-\frac{1}{\alpha^2}\nabla^2\phi_n + V\phi_n = E_n\phi_n, \tag{i}$$

and

$$-\frac{1}{\alpha^2}\nabla^2\bar{\phi}_m + V\bar{\phi}_m = E_m\bar{\phi}_m. \tag{ii}$$

Multiplying both sides of equation (i) by $\bar{\phi}_m$ and integrating over the configuration space, the result is

$$-\frac{1}{\alpha^2}\int \bar{\phi}_m\nabla^2\phi_n d\tau + \int V\bar{\phi}_m\phi_n d\tau = E_n\int \bar{\phi}_m\phi_n d\tau. \tag{iii}$$

Repeating the operation with ϕ_n on equation (ii) the result is

$$-\frac{1}{\alpha^2}\int \phi_n\nabla^2\bar{\phi}_m d\tau + \int V\phi_n\bar{\phi}_m d\tau = E_m\int \phi_n\bar{\phi}_m d\tau. \tag{iv}$$

Subtracting (iv) from (iii) we obtain the relation

$$-\frac{1}{\alpha^2}\int (\bar{\phi}_m\nabla^2\phi_n - \phi_n\nabla^2\bar{\phi}_m) d\tau = (E_n - E_m)\int \bar{\phi}_m\phi_n d\tau. \tag{v}$$

From equation (57) it follows that the left-hand side of (v) is equal to

$$\frac{1}{\alpha^2}\int\int \left(\bar{\phi}_m\frac{\partial\phi_n}{\partial n} - \phi_n\frac{\partial\bar{\phi}_m}{\partial n}\right) dS. \tag{vi}$$

Now it has been postulated that the solutions of the S. equation must represent functions which vanish at the limits of the configuration space. For instance,

the functions ϕ_n, ϕ_m, etc., must vanish for $r = \infty$, or for $x = y = z = \pm \infty$. If we take the bounding surface S at these limits, then the integral in (vi) vanishes. Consequently,

$$(E_n - E_m)\int \bar{\phi}_m \phi_n d\tau = 0,$$

and if the state corresponding to E_n is not identical with that corresponding to E_m, then $E_n - E_m \neq 0$. Hence

$$\int \bar{\phi}_m \phi_n d\tau = 0 \text{ for } n \neq m,$$
$$= N^2 \text{ for } n = m,$$

where $1/N$ is the normalizing factor.

16. Solution of Laplace's Equation in Terms of Zonal Harmonics.[5] In Chapter VI the solution of Legendre's equation (6.21)

$$(1 - x^2)\frac{d^2X}{dx^2} - 2x \cdot \frac{dX}{dx} + \alpha^2 X = 0 \tag{59}$$

was obtained by assuming that X could be expressed as a series of terms of the form $a_n x^n$ and then determining the coefficients from the condition

$$\alpha^2 = m(m + 1),$$

where $m = 0, 1, 2$, etc.

It is possible, however, to obtain a solution of equation (59) by an entirely different method which is of special interest in connection with the theory of potentials.

The potential energy due to a unit of mass (or unit of charge) concentrated at a given point $x_1 y_1 z_1$ is

$$V = \frac{1}{r} = \frac{1}{\sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2}}. \tag{i}$$

That this is a solution of Laplace's equation

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0$$

is readily verified. By differentiation of (i) we obtain

$$\frac{\partial V}{\partial x} = -\frac{(x - x_1)}{\{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2\}^{\frac{3}{2}}},$$

and

$$\frac{\partial^2 V}{\partial x^2} = \frac{3(x - x_1)^2}{\{(x - x_1)^2 + \ldots\}^{\frac{5}{2}}} - \frac{1}{\{(x - x_1)^2 + \ldots\}^{\frac{3}{2}}}.$$

From similar relations for the derivatives with respect to y and z, we obtain the result

$$\nabla^2 V = \frac{3}{\{(x - x_1)^2 + \ldots\}^{\frac{5}{2}}} - \frac{3}{\{(x - x_1)^2 + \ldots\}^{\frac{5}{2}}} = 0.$$

[5] Based on discussion by Byerly, Chapter V.

In terms of spherical coördinates r, θ, and η, (i) becomes

$$V = [r^2 - 2rr_1\{\cos\theta\cos\theta_1 + \sin\theta\sin\theta_1\cos(\eta - \eta_1)\} + r_1^2]^{-\frac{1}{2}}, \tag{ii}$$

which is a solution of Laplace's equation in spherical coördinates.

Let us consider the case in which $\theta_1 = 0$, and V is independent of η. Then

$$V = (r^2 - 2rr_1\cos\theta + r_1^2)^{-\frac{1}{2}} \tag{iii}$$

is a solution of Laplace's equation

$$\nabla^2 V = \frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\cdot\frac{\partial V}{\partial\theta}\right) = 0. \tag{iv}$$

Equation (iii) may be written in the form

$$V = \frac{1}{r}\left(1 - 2\frac{r_1}{r}\cos\theta + \frac{r_1^2}{r^2}\right)^{-\frac{1}{2}}, \tag{v}$$

or

$$V = \frac{1}{r_1}\left(1 - 2\frac{r}{r_1}\cos\theta + \frac{r^2}{r_1^2}\right)^{-\frac{1}{2}}. \tag{vi}$$

For $r_1/r < 1$, the expression in parentheses in (v) may be expanded into a convergent series, so that

$$V = \frac{1}{r}\sum p_m\left(\frac{r_1}{r}\right)^m, \tag{vii}$$

where p_m is a function of $\cos\theta$.

Hence

$$\frac{\partial V}{\partial r} = -\frac{1}{r^2}\sum p_m\left(\frac{r_1}{r}\right)^m - \frac{1}{r}\sum p_m\cdot m\cdot\frac{r_1^m}{r^{m+1}},$$

$$r^2\frac{\partial V}{\partial r} = -\sum p_m\left(\frac{r_1}{r}\right)^m(m+1),$$

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) = \sum p_m\,(m+1)m\cdot\frac{r_1^m}{r^{m+1}},$$

and

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) = \sum\frac{r_1^m}{r^{m+1}}\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dp_m}{d\theta}\right).$$

Consequently,

$$\nabla^2 V = \sum\frac{r_1^m}{r^{m+1}}\left[m(m+1)p_m + \frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dp_m}{d\theta}\right)\right] = 0.$$

Since this must be valid for all values of $r > r_1$, the coefficient of each power of r must vanish identically, and therefore

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dp_m}{d\theta}\right) + m(m+1)p_m = 0. \tag{viii}$$

Now if we set $x = \cos\theta$, then

$$\frac{dy}{d\theta} = \frac{dy}{dx}\cdot\frac{dx}{d\theta} = -\frac{dy}{dx}\cdot\sin\theta,$$

and equation (viii) assumes the form of Legendre's equation,

$$(1-x^2)\frac{d^2p_m}{dx^2} - 2x\frac{dp_m}{dx} + m(m+1)p_m = 0. \tag{60}$$

Hence, equation (vii) is a solution of this equation.

In a similar manner it can be shown that

$$V = \frac{1}{r_1}\sum p_m\left(\frac{r}{r_1}\right)^m$$

is a solution for the case $r/r_1 < 1$.

It now remains to determine the exact form of p_m. Let us consider the case $r > r_1$, and the expansion of the expression in equation (v). Let $\cos\theta = x$, and $r_1/r = z$. Then

$$(1-2zx+z^2)^{-\frac{1}{2}} = \{1-z(2x-z)\}^{-\frac{1}{2}} = 1 + \frac{z}{2}(2x-z)$$

$$+\frac{1\cdot 3}{2\cdot 4}z^2(2x-z)^2 + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}z^3(2x-z)^3 + \ldots$$

The coefficient of z^m is p_m, and the values of this coefficient are readily obtained. For the low values of m they are as follows:

$$p_0 = 1;\ p_1 = x;\ p_2 = \frac{3\cdot 1}{2!}\left(x^2 - \frac{2\cdot 1}{2\cdot 3}\right);$$

$$p_3 = \frac{5\cdot 3\cdot 1}{3!}\left(x^3 - \frac{3\cdot 2}{2\cdot 5}x\right);$$

and for z^m, the coefficient is

$$p_m = \frac{(2m-1)\ (2m-3)\ldots 1}{m!}\left[x^m - \frac{m(m-1)}{2(2m-1)}x^{m-2}\right.$$

$$\left. + \frac{m(m-1)\ (m-2)\ (m-3)}{2\cdot 4\cdot (2m-1)\ (2m-3)}x^{m-4} - \ldots\right]. \tag{61}$$

But this is the Legendre polynomial $P_m(x)$. Hence for $r_1/r < 1$,

$$V = \frac{1}{r}\left[P_0(\cos\theta) + \frac{r_1}{r}P_1(\cos\theta) + \frac{r_1^2}{r^2}P_2(\cos\theta) + \ldots\right.$$

$$\left. + \frac{r_1^m}{r^m}P_m(\cos\theta) + \ldots\right] \tag{ix}$$

is a solution of Legendre's equation (60), where m is a positive integer.

Also

$$V = r^m P_m(\cos\theta)$$

and

$$V = \frac{1}{r^{m+1}} P_m(\cos\theta)$$

are solutions of equation (iv).

17. Potential at a Point Due to a Charge Distribution. Let us consider the potential due to a homogeneous sphere. Let the radius of the sphere be a, and R the distance of the point P from the center of the sphere (see Fig. 8).

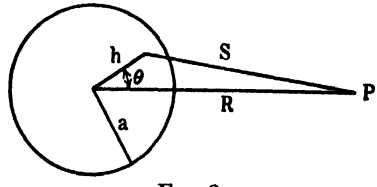


FIG. 8.

Then

$$V = \int \frac{\sigma d\tau}{S} = \int \frac{\sigma h^2}{S} \sin\theta d\theta d\eta dh,$$

where $\sigma = \sigma(r)$ is the charge distribution function and S denotes the distance from any point (inside or outside the sphere) to the point P.

Now

$$S^2 = h^2 + R^2 - 2hR\cos\theta.$$

Differentiating for a constant value of h,

$$SdS = hR\sin\theta d\theta.$$

Hence

$$V = \int \frac{\sigma h}{R} dSd\eta dh.$$

For $R > a$,

$$V = \frac{2\pi\sigma}{R}\int_0^a \int_{R-h}^{R+h} hdhdS = \frac{4\pi\sigma}{R}\int_0^a h^2 dh = \left(\frac{4\pi\sigma a^3}{3}\right)\frac{1}{R}, \tag{i}$$

since σ is independent of θ and η. That is, for an *external point the potential due to a sphere is the same as if the entire charge were concentrated at the center.*

For

$$R = a,\ V = \frac{4\pi\sigma a^2}{3}. \tag{ii}$$

Since V is constant throughout a conductor the potential at a point inside the sphere must be the same as that at the surface. That is, for a point inside a sphere of given radius a, the potential is independent of the distance from the center and given by equation (ii).

Let us now consider the potential at a point r_1 due to a charge distribution $\sigma = \sigma(r)$ which extends from $r = 0$ to $r = \infty$. It follows from equations (i) and (ii) that

for $r < r_1$,

$$V_{\mathrm{I}} = \frac{1}{r_1}\int_0^{r_1} \sigma(r) \cdot 4\pi r^2 dr, \tag{iii}$$

and for $r > r_1$,

$$V_{\mathrm{II}} = \int_{r_1}^{\infty} \frac{\sigma(r)}{r} \cdot 4\pi r^2 dr. \tag{iv}$$

Hence the potential energy due to the whole distribution is

$$V = \frac{4\pi}{r_1}\int_0^{r_1} \sigma(r) r^2 dr + 4\pi \int_{r_1}^{\infty} \sigma(r) \cdot r dr. \tag{v}$$

REFERENCES

Further remarks on the topics discussed in the previous sections will be found in some of the treatises mentioned in Appendix I and also under " Collateral Reading " in connection with Chapters II and IV, such as those by L. PAGE; SLATER and FRANK; R. A. HOUSTOUN; and JOOS.

As mentioned already, the writer has drawn freely from the work by R. E. DOHERTY and E. G. KELLER, " Mathematics of Modern Engineering," John Wiley & Sons, New York, 1936.

O. D. KELLOGG'S "Foundations of Potential Theory," Julius Springer, Berlin, 1929, is a scholarly treatise on the whole subject and treats many of the topics, especially Legendre functions and the theory of transformation of coördinates, in a very comprehensive manner.

INDEX

A

Action, 97
Action integrals, 99
Activation energy, 376*ff*
Algebraic equations, solutions of, 414
Alpha-particle, emission of, from nucleus, 71
Ammonia molecule, bond-structure of, 359, 361
Amplitude, of motion, for coupled pendulums, 212
  of wave motion, 24, 29
Amplitude function, *see* Eigenfunction
Angular momentum, of electron spin, 279
  for motion of rotator, 166
  for orbital motion of electron, 194, 206
Anharmonic linear oscillator, 345
Anti-bonding electrons, 356
Anti-parallel spin, for electrons in helium, 280
  for protons in hydrogen molecule, 352
Antisymmetrical eigenfunctions, for electron, in adjacent boxes, 215
    in box, 57
    in ionized hydrogen molecule, 218
  for excited states of helium, 271
  for hydrogen molecule, 315, 319
  for many-electron atoms, 305
  for orthohelium, 278
  for orthohydrogen, 353
  in relation to Pauli's Principle, 281
Antisymmetrical modes of vibration, for coupled dipoles, 220
  for coupled oscillators, 222
  for electron in adjacent boxes, 215
Array, *see* Matrix
Arrhenius equation, 376
Associated Laguerre functions, 185
  normalization of, 187
Associated Legendre functions, 154
  normalization of, 157
Aston, F. W., 4
Atomic energy states, Bohr theory of, 4, 101, 178
  for helium, 277
  for hydrogen, 185, 191
  for sodium, 6
Atomic orbitals, 308, 329, 359
Atomic units, 262, 309
Atomic wave functions, *see* Single-electron functions
Atoms with many electrons, Hartree's treatment of, 303
  Slater's treatment of, 304
Autoelectronic effect, 73
Average values, in classical dynamics, of r, 103
  in classical mechanics, of $1/r$, 103
  for linear oscillator, of x, p, and p^2, 125*ff*
  in quantum mechanics, 125*ff*
    of r, 200
    of r^s, 201
    of $1/r$, 202
Axes, principal (normal coördinates), 221
Azimuthal quantum numbers, 102, 178, 192

B

Band spectra for molecules, 340*ff*
Barber, W. G., 339
Beach, J. Y., 232
Bell, E. T., 2
Benzene, valence-bond structures for, 383
Bethe, H., 157, 188, 202, 207, 262, 303, 309, 339, 407, 409
Beutler, H., 335, 347, 352, 355
Bichowsky, F. R., 335
Binding electrons, 356
Birge, R. T., 208, 348, 411
Black-body radiation, distribution law for, 125

Bohr, N., 1, 2, 4, 8, 12, 17, 37, 93, 196, 388
Correspondence Principle, 17, 231, 388
theory of hydrogen atom, 89, 98
theory of origin of spectral lines, 8
Bohr orbit, mean values for, 103, 105
radius of, 102, 262
Boltzmann, L., 16
Boltzmann's distribution law, 354
Bond wave functions, 358, 361
Bonds, directed valence, 358, 361
homopolar type, 310, 318, 356
ionic type, 330*ff*
Bonhoeffer, K., 352
Born, M., 101, 103, 110, 161, 220, 234, 238
Boundary conditions, 34, 48, 56, 64, 134
Bragg, Wm., 9
Bridgman, P. W., on the nature of concepts in physics, 3
Bromine, potential energy (Morse) curve for, 380
Byerly, W. E., 110, 151, 155, 176, 405

C

Canonical form of equations of motion, 94, 99
Canonical set of structures, 382, 383
Canonically conjugated variables, 96
Carbon atom, tetrahedral eigenfunctions for, 361*ff*
Carbon monoxide molecule, bond-structure of, 361
Central force system, classical dynamics treatment of, 89*ff*
quantum mechanics treatment of, 180*ff*
Characteristic equation, 22
Characteristic function, 97
Characteristic functions, *see* Eigenfunctions
Characteristic values, *see* Energy values
Charge distribution, potential at a point due to, 266, 436
Chemical bonds, *see* Bonds
Classical atomic dynamics, 77*ff*
relation to quantum mechanics, 398*ff*
Clusius, K., 352
Cold cathodes, electron emission from, 73
Commutative operators, 43
Complex conjugate quantities, 23, 415
Compton, A. H., 1, 10, 11, 17
Compton effect, 10
Condon, E. U., 53, 68*ff*, 76, 122, 136, 155, 156, 158, 186, 240, 244, 260*ff*, 288, 316, 345, 407, 409
Confocal elliptic coördinates, 336*ff*
Conservative system, 96
Constant *a* of van der Waals' equation, derivation of, 233
interpretation of, 208, 210
Constant *b* of van der Waals' equation, 208, 210, 235
Contact transformation, 96
Continuity, equation of, 423
Coolidge, A. S., 333*ff*, 356, 381
Coördinates, confocal elliptic, 336*ff*
cylindrical, 82, 84
elliptical, 294
generalized, 82
normal, 221*ff*
orthogonal curvilinear, 84, 138, 430
polar, 106
rules for transformation of, 106, 138, 336, 429*ff*
spherical, 83, 138
Copeland, L. C., 335
Correlation energy, 299, 329
Correspondence, Principle of, 17, 231, 388
Coulomb integral (energy), 272, 275, 307, 317*ff*, 356*ff*, 370*ff*
Coupled electric dipoles, 220*ff*
Coupled linear oscillators, 225*ff*
Courant, R., 46, 116, 138, 155, 177
Critical constants (van der Waals), 208, 210, 233
Critical potentials, 4, 7
Cross product of two vectors, 420, 426
Crystal potential constant, 236
Curl operator, 426
Current, expression for, in quantum mechanics, 402
Cylindrical coördinates, 82, 84

D

d'Alembert's principle, 80, 283
Darrow, K. K., 17, 18, 31, 136, 409
Darwin, C. G., 17, 211, 238, 275, 407
Davisson, C. J., 11, 18

de Boer, J. H., 238
de Broglie, L., 11, 36, 41, 389
de Broglie wavelength, 11, 37, 412
Debye, P., 231, 238
Definite integrals, table of, 415
Degeneracy due to orbital-spin interaction, 306, 307
Degenerate energy states, for helium atom, 269
for hydrogen atom, 103, 181, 191
for hydrogen molecule, 312
perturbation theory treatment of, 249*ff*
for rigid rotator, with fixed axis, 146
with free axis, 156
Degrees of freedom, 85
del operator, 423
del squared operator, 423
Dempster, A. J., 11
Dennison, D. M., 352, 355
Derivative of a vector, 422
Determinantal equation, 252, 270, 327
Determinants, 248, 414
Determinant-type wave functions, 305, 365*ff*
Diatomic molecule, dissociation energy for, 345
energy levels for, electronic, 341
rotational, 349
vibrational, 343
Differential equation (*see also* S. equation), canonical, 95
Fick, 427
for heat flow, 427
for hyperbolic functions, 27
Lagrange, 88
Laguerre, 186
Laplace, 429
Legendre, 148
for motion under constant force, 20
for orbit of electron in Bohr atom, 91
Poisson, 429
for propagation of wave motion, 29, 427
significance of, 20
for simple harmonic motion, 22
for vibration of a string, 31
Diffusion, equation for, 427
of particles in zero field, 390, 405
Dipole interaction, 220
Dipole moment, classical theory of radiation for, 385
electric, definition of, 219
as vector quantity, quantum mechanics interpretation of, 386, 392*ff*
Dirac, P. A. M., 1, 19, 77, 285, 409
Directed valence bonds, Pauling's theory of, 358*ff*
Discriminant, 85, 139, 431
Disintegration, radioactive, 70
Dispersion forces, 238, 259, 260
Dissociation energy, of diatomic molecules, 345
of hydrogen molecule, 332
Distance, element of, in coordinates, curvilinear, 431
spherical, 84
Distribution Law, Maxwell-Boltzmann, for velocities, 125
Planck, for black-body radiation, 125
Divergence operator, 423
Doherty, R. E., 46, 419, 437
Dot product of vectors, 420
Dushman, S., 17, 73, 75, 125, 196, 259, 318, 329, 352, 358

E

Eccentric anomaly, 108
Eckart, C., 75, 285, 290
Eddington, A. S., 13, 19
Eigenfunctions, antisymmetrical, *see* Antisymmetrical Eigenfunctions
complete, including spin, for helium atom, 281
for many-electron atom, 304
for electron in box, 57
electron spin, 281, 319
expansion of, in terms of an orthogonal system of functions, 171*ff*, 242, 245
geometrical interpretation of, for rigid rotator, 158*ff*
for linear harmonic oscillator, 116, 120
molecular orbital, 330*ff*, 382
for motion of electron in zero field, 39
normalization factor for, *see* Normalizing factor
normalized, for coupled oscillators, 227

Eigenfunctions, normalized, for coupled oscillators, first-order perturbation term approximation, 256
normalized form of, definition of, 60, 117
normalized spherical, for hydrogen atom, 190
orthogonality of, 62, 121, 133, 153
orthonormal, definition of, 63
for polyatomic molecules, Slater's treatment of, 364*ff*
proton spin, 353
quantum mechanics interpretation of, 192
for reflection and transmission of electrons at potential barriers, 49, 52, 64, 66
relation between, and particle velocity, 402*ff*
for rigid rotator, with fixed axis, 145
with free axis, 150, 155
rotational, for ortho- and para-hydrogen, 353
single-electron, 308
for solution of radial S. equation, 186, 188, 189
for string fastened at both ends, 36
symmetrical, *see* Antisymmetrical Eigenfunctions
valence bond, *s*- and *p*-, 359
tetrahedral *sp*, 361*ff*
zero-order, for helium atom, 271
for hydrogen molecule, 313
Eigenvalue, definition of, 36
Eigenvalues, *see* Energy values
Einstein, A., 1, 8
photoelectric effect, law of, 8
Eisenschitz, R., 232, 257, 258, 261, 417
Eldridge, J. A., 5
Electric dipole, moment of, 219*ff*
Electric oscillator, radiation from, 385*ff*
Electron density distribution, *see* Probability distribution functions
Electron-pair bond, 310, 318, 356
Electron spin, 278*ff*
magnetic moment of, 279
quantum number, 279
Electrons, anti-bonding, 356
binding, 356
in a box, 55*ff*
Electrons, diffraction of, 11
emission of, in intense fields, 72
from a monatomic film, 75
exclusion principle for, 279
motion of, in field-free region, 39
promotion of, 362
reflection and transmission of, through potential barriers, 47*ff*
Ellipse, eccentric anomaly, 108
eccentricity of, 92
equation for, in polar coördinates, 106*ff*
Elliptic orbits, for central force systems, 89*ff*
mean value, of radius, 103
of reciprocal radius, 105
Elliptical coördinates, 394
Emde, F., 177, 325, 418
Emission of radiation, *see* Radiation
Energy, of activation, 376*ff*
dissociation, of diatomic molecules, 345
of H_2, 332
evaporation, 236
Hamiltonian function for total, 42, 96, 111, 166, 180, 226
interaction, of three univalent atoms, 364, 374
of four univalent atoms, 376
kinetic, definition of, 78
perturbation, degenerate case, first-order, 249*ff*
for hydrogen molecule, first-order, Heitler-London treatment of, 310
interpretation of, 315
non-degenerate case, first-order, 241*ff*
second-order, 245*ff*
potential, definition of, 79
Morse equation for, of diatomic molecule, 346
resonance, in molecules, 381*ff*
of sublimation, 236
total, definition of, 80
van der Waals interaction, for hydrogen atoms, 228*ff*
for other atoms and molecules, 233
Energy levels, spectral, of helium, 277
of sodium, 6

Energy value for normal state of helium atom, 268, 292, 299
Energy values, for coupled harmonic oscillators, 223, 226, 253
discrete, for atomic systems, 4
for electron in a box, 58
for excited states of helium atom, 271
for hydrogen atom, 191
quantum mechanics treatment of, 185
Wilson-Sommerfeld treatment of, 101
for hydrogen molecule, Heitler-London treatment of, 315
variational treatment of, 326
for linear harmonic oscillator, 115
for rigid rotator, with fixed axis, 145
with free axis, 149, 156
for string fastened at both ends, 36
Equation, of continuity, 423
for wave motion, 32, 427
Equations of motion, in Hamiltonian form, 95*ff*
in Lagrangian form, 88*ff*
Error function (erf) of Gauss, 136, 416, 418
Eucken, A., 352
Euler's constant, 325, 416
Euler's equation, 283
Euler's integral of the second kind, 417
Euler's theorem for homogeneous function, 86
Evaporation, energy of, 236
Even and odd functions, 57, 59, 151
Exchange force, 276
Exchange integral (energy), 272, 274, 307, 317*ff*, 324, 356*ff*, 370*ff*
Exclusion principle of Pauli, 279, 282, 305, 319
Expansion of a function in series of orthogonal functions, 171*ff*, *see also* Series developments
Exponential functions, expression for, in terms of hyperbolic functions, 28
in terms of trigonometric functions, 23
Exponential integral, 325, 416
Eyring, H., 357, 358, 365, 376, 378, 379, 381, 382

F

Farkas, A., 352, 380
Fick's equation for diffusion, 427
Fock, V., 304, 307
Force, between dipoles, 227
definition of, 78
quantum mechanics concept of, 276
Forces, dispersion, 238, 259*ff*
van der Waals, 208*ff*
Fowle, F. E., 208, 418
Fowler, R. H., 73*ff*, 236, 238
Franck, J. 4*ff*
Frank, N. H., 81, 85, 110, 140, 207, 238, 261, 407
Frank, P., 46, 116, 410
Free particle, in a box, 55*ff*
wave mechanics representation of motion of, 39*ff*
Frenkel, J., 55, 76, 168, 407, 409
Frequencies of modes of vibration for coupled oscillators, 222*ff*
Frequency, of harmonic oscillator, 22, 111
of vibration in molecules, 340*ff*
Functions (*see also* Eigenfunctions), exponential, 22, 28, 34, 39, 145, 415
Hermitian, 116*ff*
hyperbolic, 28, 216, 414*ff*
hypergeometric, 75
Laguerre, 182*ff*
Legendre, 150*ff*
trigonometric, 23, 34, 57, 63, 145, 171, 415
Fundamental differential equation, for hyperbolic functions, 27*ff*
for particle under constant force, 20
for simple harmonic motion, 22*ff*

G

Gamma function, 417*ff*
Gamma-ray microscope, 14
Gamow, G., 70
Gauss's error function, 136, 416, 418
Gauss's law, 428
Gauss's theorem, 424
General solution of differential equation, 22, 28
Generalized coördinates (*see also* Coordinates), 82*ff*

Generalized momenta, 94
Generating function, for associated Laguerre polynomials, 185
for associated Legendre polynomials, 155
for zonal harmonics, 433
Germer, L. H., 11, 18
Gershinowitz, H., 376
Goppert-Mayer, M., 220
Goudsmit, S., 196, 279, 282, 410
Gradient, 422
Gravity, motion under, 80, 407
Green's theorem, 295, 431
Group velocity, relation to particle velocity, 402*ff*
relation to phase velocity, 398*ff*
Gurney, R. W., 17, 68*ff*, 76, 214, 407, 409

H

Haas, A., 409
Half-integral quantum numbers, 115
Hamiltonian function, for coupled oscillators, 221
definition of, 96
for harmonic oscillator, 111
for hydrogen-like atom, 99
for rigid rotator, 161
in spherical coördinates, 140
Hamiltonian form of function for total energy, 42, 96, 111, 166, 180, 221
Hamiltonian operator, 203, 284, 290, 294, 309, 321, 365*ff*
Hamilton's function, 97
Hamilton's principle, 96, 283
Harmonic oscillator, average value, of p for, 128
of x for, 127
classical treatment of, 22*ff*
distribution functions for, 122
eigenfunctions for, 116, 120
energy levels for, 115
probabilities of transition for, 131*ff*
quantum mechanics treatment of, 111*ff*
radiation from, 385*ff*
selection rule for, 393
Harmonic oscillators, coupled, 219*ff*
Harmonics, surface spherical, 158*ff*
Harteck, P., 352
Hartree, D. R., 262, 303, 304
Hassé, H. R., 230, 232, 339
Haurwitz, E. S., 308
Heat flow, equation for, 423, 426
Heisenberg, W., 1, 12, 13, 15, 19, 131, 352, 398, 409
Heitler, W., 257, 310, 315, 316, 318, 325, 329, 356, 358, 370
Heitler-London type of bond, 310, 318, 356
Helium atom, ionization potential of, 268, 293, 300
multiplet levels in spectrum of, 280
perturbation method, for excited states of, 268*ff*
for normal states of, 265*ff*
spectroscopic energy levels for, 277
variational method for normal state of, 289, 293
Helium atoms, van der Waals interaction of, 233
Hermite, C., 116
Hermitian matrix, 249
Hermitian polynomials, 116*ff*
Hertz, G., 4, 5
Hertzian oscillator, 385
Herzfeld, K. F., 238
Hilbert, D., 46, 116, 138, 155, 177
Hill, E. L., 409
Hiller, K., 352
Hitchcock, F. L., 46
Homopolar bond, 310, 318, 356
Homopolar terms in bond functions, 331
Houston, W. V., 46
Houstoun, R. A., 46, 176
Hückel, E., 358, 383, 384
Humphrey, D., 46, 176
Hund, F., 330, 352, 409
Hydrogen atoms, elastic reflection of, 318
interaction of, with para- or ortho-hydrogen, 380
molecule formation from, 310*ff*
van der Waals interaction of, 228
Hydrogen bromide, potential energy (Morse) curve for, 380
Hydrogen-like atom, angular momentum of electron in, 206
average values, of r for, 200
of $1/r$ for, 202
classical treatment of, 99*ff*
distribution functions for, angular, 193
radial, 195

Hydrogen-like atom, energy states of, 101, 185, 187, 263
normalized functions for, radial, 188
spherical, 190
orbital radius of, 178
radial equation for, 182
S. differential equation for, 180
Hydrogen molecule, dissociation energy of, 332
energy states of, 341
rotational, 349
vibrational, 343
excited states of, 342, 348, 351
potential energy (Morse) curve for, 341, 346
theory of formation of, Heitler-London method for, 310*ff*
method of molecular orbitals for, 329
variational method for, 326
Hydrogen molecule ion, plots of eigenfunctions for, 218
Hydrogen molecules, repulsive interaction of, 357
Hylleraas, E. A., 293, 299, 300, 332, 333
Hyperbolic functions, 28, 66, 216, 414, 415
Hypergeometric type of S. equation, 75
Hypotheses in physics, rôle of, 2

I

Indeterminacy, Principle of, 12, 40, 44, 124, 205, 302, 398
Ingersoll, L. R., 176
Ingham, A. E., 236
Inhomogeneous differential equation, 242, 269, 314
Integral invariants, 98
Integral logarithm, 325, 416
Integrals, Coulomb type, *see* Coulomb integral
definite, expressions for, 415*ff*
exchange type, *see* Exchange integral
Interaction, of two hydrogen atoms, elastic reflection in, 318*ff*
molecule formation in, 318*ff*
of two pendulums, 211
(repulsive) of hydrogen molecules, 357
van der Waals, 208*ff*
of helium atoms, 233
of hydrogen atoms, 228, 256
of two linear oscillators, 225, 252
Interatomic distance in hydrogen molecule from band spectrum, 347
Ionic bonds, 331
Ionic terms in bond functions, 330
Ionization energy, of hydrogen atom, 412
of helium atom, 268

J

Jacobian, 294, 339
Jahnke, E., 177, 325, 418
James, H. M., 333*ff*, 356, 381
Jeans, J. H., 176, 230, 257
Jeppesen, C. R., 342, 346*ff*, 351, 355
Jones, J. E., 236
Joos, G., 110, 410

K

Kekulé structures, resonance of, 383
Keller, E. G., 46, 419, 437
Kellogg, O. D., 176, 437
Kelvin, Lord, 1
Kemble, E. C., 168, 177, 240, 261, 288, 409
Kennard, E. H., 17, 407
Kepler, third law of planetary motion, 93
Kimball, G. E., 382
Kinetic energy, average value of, for hydrogen atom, 202
definition of, 78
expression for, in terms of gradient of eigenfunction, 204, 287, 296, 301
relation to solution of S. equation, 201*ff*, 301
virial theorem for, 109
Kinetic-potential, *see* Lagrangian function
Kirkwood, J. G., 210, 232*ff*
Kistiakowsky, G. B., 384
Kronig, R. de L., 343 355

L

Lagrange, J. L., 80, 86
Lagrange's equations of motion, 88, 283
Lagrangian function, 88, 283
Laguerre polynomials (functions), 185*ff*
Landé, A., 407
Langmuir, I., 2
Laplace's equation, 429
solution of, in zonal harmonics, 433

Laplacian operator, 423
in coördinates, Cartesian, 137
confocal elliptic, 339
orthogonal curvilinear, 139
spherical, 139, 141, 180
Laue, M. v., 182, 409
Least action, principle of, 98
Legendre polynomials (functions), associated, 154*ff*, 418
of order zero, 150*ff*, 435
Legendre transformation, 95
Legendre's equation, 148, 435
Lennard-Jones, J. E., 224, 228, 230, 232, 234*ff*, 257, 307
Lindsay, R. B., 17, 86, 89, 110, 261, 410
Linear harmonic oscillator, *see* Harmonic oscillator
Line integrals, 421
relation to surface integrals, 426
Localized pairs, method of, 374
London, F., 209, 210, 219, 230, 232, 233, 236, 257, 258, 261, 310, 315, 316, 318, 325, 356, 358, 370, 376, 377, 417

M

MacRobert, T. M., 155, 176
Madelung, E., 46
Magnetic moment, of Bohr magneton, 412
of electron spin, 279
Magnetic quantum number, 192
Magneton, Bohr, 412
Main-Smith, J. D., 280
Many-electron atom, method of Hartree, 301
method of Morse, Young and Haurwitz, 308
method of Slater, 304
Margenau, H., 17, 86, 89, 110, 232, 261, 410
Mathematical formulas, special tables of, 413
Matrix, definition of, 248
Hermitian, 249
Matrix elements, calculation of, for helium atom, 272
for hydrogen molecule, 327
in treatment of polyatomic molecules, 365*ff*
Maxwell, Clerk, 16, 385
Maxwell-Boltzmann distribution law for velocities, 125
Mayer, J. E., 234
Mean values, *see* Average values
Mellor, J. W., 46, 399
Methane molecule, bond structure of, 362
Michelson, A. A., 1
Minimum energy, principle of, 283
Mises, R. v., 46, 116, 410
Molecular orbitals, 329, 333
Molecules, energy levels for, *see* Hydrogen molecule
Momenta, generalized, definition of, 94
Momentum, angular, for electron spin, 279
for motion of rotator, 166*ff*
unit of, 412
Momentum operator, 167
Morley, E. W., 1
Morse, P. M., 122, 136, 155, 156, 158, 186, 240, 244, 260, 261, 288, 308, 316, 345, 346, 407, 409
Morse (potential energy) function for diatomic molecules, 346, 379
Moseley, J. H., 4
Motion under action of constant force, 81
Mott, F. N., 41, 42, 61, 76, 179, 182, 212, 218, 303, 409
Muller, A., 238
Mulliken, R. S., 330
Multiplet levels in spectrum of helium, 277, 280

N

Newton's second law of motion, 77, 402
Non-polar bond, 319, 331
Nordheim, L., 73*ff*
Normal coördinates, 221
Normal modes of vibration, for coupled oscillators, 222
for coupled pendulums, 211
Normalizing factor, definition of, 60, 117
for eigenfunctions for linear harmonic oscillator, 120
for Legendre polynomials, associated, 157
of zero order, 153
for radial eigenfunctions of hydrogen-like atom, 187

Normalizing factor, for tesseral harmonics, 157
for trigonometric functions, 63
Nuclear spin for hydrogen, 352
Nuclei, radioactive distintegration of, 70
Nullpunktsenergie, 116

O

Odd and even functions, 57, 59, 151
Operator, angular momentum, 167
commutative and non-commutative, 43, 139
del, definition of, 423
del squared, 423
divergence, 423
gradient, 422
Hamiltonian, 203, 284, 290, 294, 301, 309, 326
Laplacian, 137, 139, 423
linear momentum, 42
method of deriving S. equation, 42
vector, 419*ff*
Oppenheimer, J. R., 73
Orbital quantum numbers, 280
Orbitals, molecular, 329, 333
Orbits, for central force systems, 89*ff*
types of, on basis of Wilson-Sommerfeld conditions, 102
Orthogonal curvilinear coördinate systems, 84, 138, 430
Orthogonal functions, expansion in terms of, 171*ff*
Orthogonality of solutions of S. equation, 62, 121, 133, 153, 432
Orthohelium, 277
Orthohydrogen, 352
interaction of, with atomic hydrogen, 380
Orthonormal functions, definition of, 63
Oscillator, *see* Harmonic oscillator
anharmonic type, 345
electric, radiation from, 385*ff*
Osgood, W. F., 110

P

p-eigenfunctions, 360
Packets, wave, 398*ff*
relation to particle velocity, 402
spreading of, in time, 405
Page, L., 85, 101, 110
Parahydrogen, 352
interaction of, with hydrogen atoms, 380
Parhelium, 277
Particle, in a box, 55
in field-free region, 39
Paschen-Back effect, 279
Pauli exclusion principle, 279
application, to helium atom, 280
to hydrogen molecule, 319
in relation to antisymmetrical functions, 282, 305
Pauling, L., 59, 76, 85, 94, 110, 117, 136, 138, 155, 163, 177, 188, 191, 195, 196, 201, 207, 232, 238, 258, 261, 275, 282, 293, 304, 306, 307, 309, 331, 349, 350, 355, 358, 359, 361, 362, 364, 374, 381*ff*, 392, 409, 410
Peirce, B. D., 418
Pekeris, L. C., 349, 350
Penney, W. G., 329, 330, 357, 374*ff*
Perturbation method, application of, to coupled oscillators, 253
to helium, excited state, 269
normal state, 265
to hydrogen molecule, 314
to Stark effect, 259
to van der Waals interaction of two hydrogen atoms, 256
to Zeeman effect, 260
Perturbation theory, first-order terms, for degenerate states, 249
for non-degenerate states, 241
second-order terms, 245
Phase velocity, 33, 41, 401
Phillips, H. B., 46
Photoelectric effect, 8
Photons, 9
Physical constants, values of, 411
Planck, M., 1
Planck's constant, 5, 99
Planck's distribution law, 125
Poincaré, H., 2
Poisson's equation, 429
Polanyi, M., 238, 365, 376, 378, 381
Polar bond, 331
Polar coordinates, equation of ellipse in, 106
spherical, 83

Polarizability, definition of, 220
relation, to dispersion spectrum, 237, 259, 260
to force constant for harmonic oscillator, 221
Polarization of emitted light, 394, 396
Polyatomic molecules, Slater's treatment of, 364*ff*
Polynomials, Hermitian, 116
Laguerre, 185
Legendre, 150, 157
Potential due to charge distribution, 266, 436
Potential barrier problems, 47*ff*
Potential energy, average value of, for central orbit, 105, 109
definition of, 79
for harmonic oscillator, 26, 111
for hydrogen atom, 202
in relation to solution of S. equation, 203*ff*
sign of, 81
virial theorem for, 109
Potential energy (Morse) curve for hydrogen molecule, 341
Potential energy curves, for hydrogen molecule, hydrogen bromide and bromine, 380
for van der Waals interaction between atoms, 235
Potential energy function, relation of, to solution of S. equation, 201*ff*
Power-series method for solving S. equation, 114, 148, 183
Principle, of Correspondence, 17, 231, 388
of Hamilton, 96
of Indeterminacy, 12, 40, 44, 124, 205, 302, 398
of least action, 98
of minimum energy, 283
of superposition of wave motions, 25, 216
Probability distribution functions, for electron in box, 59
for electrons in Heitler-London interaction of two hydrogen atoms, 318
for hydrogen atom, angular, 192
radial, 194*ff*
Probability distribution functions, for interaction of two harmonic oscillators, 228
for linear harmonic oscillator, 122
for motion of electron in zero field, 40
in relation to average kinetic energy, 203*ff*, 300*ff*
in relation to average potential energy, 203*ff*, 300*ff*
in relation to behavior of wave packets, 402, 405
for rigid rotator, with fixed axis, 145
with free axis (zonal harmonics), 162*ff*
Probability, total, of occurrence of electrons in non-classical region, 53
of penetration of electron through finite barrier, 65, 67
Promotion of bond eigenfunctions, 362

Q

Quadratic homogeneous functions, 86
Quadrupole interaction, 221, 232
Quantum condition for Bohr orbits, 37
Quantum conditions of Wilson-Sommerfeld, 98*ff*
Quantum effects, survey of, 4*ff*
Quantum number, for angular momentum, of orbital motion of electron, 206
of rotation, 166*ff*
of spin, 279
azimuthal, 102, 178, 181, 192
for electron spin, 279
half integral, 115
magnetic, 192
orbital, 279
radial, 189
in relation to type of orbit, 102
rotational, 145, 149, 181, 341*ff*
total, 101, 178, 192
vibrational, 115, 341*ff*
Quantum numbers, 99
and Pauli Exclusion Principle, 279
in relation to spectral terms, 192
significance of, 189

R

Radial distribution function for hydrogen atom, 194*ff*

Radial equation for hydrogen atom, 181*ff*, 263
Radial quantum number, 189
Radiation, electric oscillator, intensities of lines in, 393*ff*
emission of, classical theory of, 385*ff*
quantum mechanics theory of, 392*ff*
Planck's law of, 125
selection rule for, from rotator, 394*ff*
Radioactive disintegration, 70
Rate of chemical reaction, temperature coefficient of, 376
Recursion formulas, for Hermite polynomials, 177
for Legendre polynomials, 397
Reduced mass, 94, 144, 344
Reflection coefficient for electrons at semi-infinite potential barrier, 50
Resonance among bond structures, 381*ff*
Resonance integral (energy), *see* Exchange integral
Resonance radiation, 7
Richtmyer, F. K., 17, 259
Ritz (variational) method, 283*ff*
Robinson, C. S., 46
Rodrigues' formula, 153
Rosen, N., 329, 332*ff*
Rotational energy levels for hydrogen molecule, 343
Rotational energy states of diatomic molecules, 341*ff*
Rotational quantum number, 145, 149, 341*ff*
Rotator, angular momentum of, 166
rigid, with fixed axis, 143*ff*
with free axis, 146*ff*
Ruark, A. E., 17, 94, 191, 355, 407, 409
Rumer, G., 358, 381
Rutherford, E., 4
Rydberg constant, 7, 101, 178, 185, 412

S

s-eigenfunctions, 359
Scalar product of vectors, 420
Scalars, 420
Schroedinger, E., 1, 11, 19, 31, 36, 37, 177, 240, 260, 284
Schroedinger equation, corresponds to Euler equation in variational problem, 284, 287
derivation of, for one coordinate, 36*ff*
for emission of electrons in intense fields, 73
for helium atom, 239, 264
for interaction of two linear oscillators, 225
involving time, 388*ff*
for linear harmonic oscillator, 111
in electric field, 225
for motion of electrons in zero field, 39
operator method for derivation of, 42
in orthogonal curvilinear coordinates, 139*ff*
orthogonality of solutions of, 62, 121, 133, 153, 432
radial, for hydrogen-like atom, 181, 263
in rectangular coordinates, in three dimensions, 137
for reflection of electrons at semi-infinite potential barriers, 48, 52
for rigid rotator, 144
in spherical polar coordinates, 180
for transmission of electrons through finite potential barrier, 63
for vibrational-rotational states of diatomic molecules, 344
Screening constant, 290
Second-order perturbation theory, 245*ff*
Sectorial harmonics, 160
Secular equation, 252, 270, 327
Seitz, F., 201, 285, 299, 318, 329
Selection rules, for linear harmonic oscillator, 393
for quantum numbers m and l, 394*ff*
Self-consistent field, method of, 303*ff*
Separation of variables in S. equation, 329
Series, spectral, of helium, 277
of sodium, 6
Series developments, for associated Laguerre functions, 186
for binomial expressions, 223, 413
for exponential functions, 22, 413
for Hermite functions, 116
for hyperbolic functions, 28, 414, 415

Series developments, for integral logarithm, 325
for Legendre functions, 150
for trigonometric functions, 23, 413
Shared electron-pair, *see* Homopolar bond
Sherman, A., 320, 330, 331, 358, 361, 362, 374*ff*
Sherman, J., 238
Shortley, G. H., 262, 409
Silberstein, L., Mathematical Tables, 118, 136, 151, 177, 414, 418
Single-electron wave functions, 308, 359
Singlet state of H_2, 319, 341
Slater, J. C., 81, 85, 110, 140, 207, 210, 232*ff*, 238, 261, 303*ff*, 307, 309, 358, 364, 366, 368, 370, 371, 381, 410
Smallwood, H. M., 238
Smyth, C. P., 238
Solution of Laplace's equation in zonal harmonics, 433
Solutions of S. equation, orthogonality of, 62, 121, 133, 153, 432
Sommerfeld, A., 99, 101, 136, 161, 177, 189, 191, 280, 409
Spatial degeneracy, *see* Degenerate energy states
Spectral terms, *see* Energy levels
Spectroscopic nomenclature for single electron states, 188, 192
Spherical harmonics, 158
Spherical polar coördinates, 83
Spin, electron, 278
angular momentum of, 279
functions, 280, 304, 306, 319, 366*ff*
of proton, 352
Spin degeneracy, 280, 306
Sponer, H., 348, 355
Stark effect, application of perturbation theory to, 259
Stokes's theorem, 426
Stoner, E. C., 280
Sublimation, energy of, 236
Sugiura, Y., 316, 317, 325, 334
Sun, C. E., 376
Superposition, of wave functions, 390
of wave motions, 25, 216
Surface harmonic wave functions, selection rules for, 394*ff*
Surface harmonics, geometrical interpretation of, 158
Surface integrals, 421
relation to line integrals, 426
relation to volume integrals, 295, 431
Symmetrical wave functions, *see* Antisymmetrical eigenfunctions

T

Taylor, H. S., 259, 387
Tesseral harmonics, 156*ff*, 181
Tetrahedral carbon atom, bond-wave functions for, 362
Theorem, of Euler, 86
of Gauss, 424
of Green, 295, 431
of Stokes, 426
Thomson, G. P., 11, 18
Tolman, R. C., 110, 125
Topley, B., 376
Total energy, definition of, 80
Total quantum number, 101, 178
Transformation of coördinates, 83, 429
generalized theory of, 429
from rectangular, to confocal elliptic, 336
to cylindrical, 82
to spherical polar, 83
Transformation of Legendre, 95
Transition probabilities, for linear harmonic oscillator, 131
for rigid rotator, 394*ff*
for rotational levels in hydrogen, 351, 353
for states corresponding to symmetrical and antisymmetrical functions, 282
Transmission coefficient for electrons, through finite potential barrier, 65, 67
through semi-infinite barrier, 50

U

Uhlenbeck, G. E., 279
Uncertainty, Principle of, *see* Indeterminacy, Principle of
Unsöld, A., 266, 267, 274, 282
Urey, H. C., 17, 94, 191, 197, 355, 407, 409

V

Valence bond eigenfunctions, 358, 361, 364, 371
Valence bonds, Heitler-London treatment of, 310*ff*
homopolar, 356
Pauling's theory of directed, 358
representation of, for carbon, 361
Slater's treatment of, 364
Van der Waals' equation, 208
Van der Waals forces, 208, 231
Van der Waals interaction energy, for hydrogen atoms, 228
for other atoms, 233
Van Vleck, J. H., 101, 109, 110, 320, 330, 331, 358, 361, 362, 374*ff*
Variation theorem, 283
Variational method, application of, to helium, 289
method of Hylleraas, 293
to other atoms, method of Morse, Young and Haurwitz, 308
to hydrogen molecule, 326*ff*
to linear harmonic oscillator, 288
Vector product, 420, 426
Vector quantity, angular momentum as, 168
electric moment as, 386, 396
Vectors, 419*ff*
Velocity of wave motion, group, 401
phase 33, 401, 427
Vibration, frequency of, for harmonic oscillator, 22, 111
modes of, for coupled oscillators, 222
for coupled pendulums, 211
for hydrogen molecule, 341
symmetric and antisymmetric modes of, for electron in two adjacent boxes, 215
Vibrational energy states of diatomic molecules, 340*ff*
Vibrational quantum number, 115, 340, 342, 346, 351
Vibrations, superposition of, 212
Virial theorem, 109
Volume element in coördinates, orthogonal curvilinear, 85, 431
spherical, 84

W

Waller, I., 201
Wang, S. C., 210, 332, 334
Water, molecule of, bond-structure of, 359, 361
Wave equation, *see* Schroedinger equation
Wave function, *see* Eigenfunction
Wavelength, de Broglie, 11, 37, 412
Wave motion, amplitude of, 24, 29
equation for propagation of, 29, 427
velocity of, group, 402
phase, 31, 401
Wave number, definition of, 7
Wave numbers, for levels in spectrum, of helium, 277
of sodium, 6
for vibrational levels in band spectrum of hydrogen molecule, 346, 348
Wave packets, 398
diffusion of, in time, 405
in relation to corpuscular motion, 402
Waves, superposition of, 25, 213
Webster, A. G., 80, 85, 110
Weinbaum, S., 332
Wheeler, A., 376
Wheland, G. W., 384
White, H. E., 188, 193, 195, 201, 207, 282, 410
Whittaker, E. T., 110
Wigner, E., 299
Wilson, E. B., Jr., 59, 76, 85, 94, 110, 117, 136, 138, 155, 163, 177, 188, 191, 195, 201, 207, 232, 238, 258, 261, 275, 282, 293, 304, 306, 307, 309, 331, 349, 350, 355, 381, 392, 409
Wilson, W., 110
Wilson-Sommerfeld rules of quantization, 99

Y

Young, L. A., 308

Z

Zeeman effect, application of perturbation theory to, 260

Zero-order eigenfunction, for helium atom, 271
for hydrogen molecule, 313
Zero-order eigenfunctions in perturbation theory, 341*ff*
Zero-order Legendre functions, 150
Zero-point energy, for harmonic oscillator, 116
for vibrational state of hydrogen molecule, 347
Zobel, O. J., 176
Zonal harmonics, 150*ff*, 159, 418